University of Alberta

Characterization of genes on human chromosome 14q13: Focus on holoprosencephaly

By

Deepak Kamnasaran

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment
of the requirements for the degree of Doctor of Philosophy

in

Medical Sciences – Medical Genetics

Edmonton, Alberta

Spring 2003

National Library
of Canada

Acquisitions and
Bibliographic Services

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale
du Canada

Acquisisitons et
services bibliographiques

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-612-82125-0
Our file Notre référence
ISBN: 0-612-82125-0

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou aturement reproduits sans son autorisation.

Canada

UNIVERSITY OF ALBERTA

Library Release Form

Name of Author: Deepak Kamnasaran

Title of Thesis: Characterization of genes on human chromosome 14q: Focus on holoprosencephaly

Degree: Doctor of Philosophy

Year this Degree Granted: 2003

Permission is hereby granted to the University of Alberta Library to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly or scientific research purposes only.

The author reserves all other publication and other rights in association with the copyright in the thesis, and except as herein before provided, neither the thesis nor any substantial portion thereof may be printed or otherwise reproduced in any material form whatever without the author's prior written permission.

DKam

3615-106 Street
Edmonton, Alberta
Canada
T6G 2H7

Date: Mar 31st / 2003

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled "Characterization of genes on human chromosome 14q13: Focus on holoprosencephaly" submitted by Deepak Kamnasaran in partial fulfillment for the degree of Doctor of Philosophy in Medical Sciences – Medical Genetics.

Diane W. Cox
Dr. D.W. Cox

Dr. M.A. Walter

J Bamforth
Dr. J.S. Bamforth

Dr. D.A. Underhill

John B Bell
Dr. J. Bell

S Scherer
Dr. S.W. Scherer

Date: 3/28/03

This thesis is dedicated to my mom and dad for their tremendous support, love and care during my graduate years.

Abstract

This dissertation describes the characterization of a holoprosencephaly locus on human chromosome 14q, designated HPE8, with the ultimate objective to identify and characterize one or more candidate genes for this disease. Holoprosencephaly is the most common congenital defect in the development of the human forebrain and associated facial structures. A spectrum of mild to severe forebrain and facial anomalies occur in this disorder. A panel of patients, who are carriers of proximal chromosome 14q rearrangements, was collected through worldwide collaborations. These were used for mapping to define an approximately 2 Mb minimal critical region for holoprosencephaly on chromosome 14q13. Both YAC and BAC contigs were built by standard methods, including chromosome walking. With the availability of human draft genomic sequence from the Human Genome Project and Celera Genomics, the sequence was annotated for 34 putative genes using bioinformatic approaches. Three genes (*C14ORF11*, *NPAS3*, *SNX6*), expressed in human fetal brain, were selected for characterization. These three candidate genes were characterized in humans in order to obtain insight into the biological function. These included the isolation of the human cDNA, putative properties using extensive bioinformatic analyses, and determination of the genomic structure and expression profile and cellular location of the protein. The murine orthologues of these three candidate genes were also studied for the temporal and spatial expression profile in developing mouse embryos. *C14ORF11*, *SNX6* and *NPAS3* are suggested to be excellent candidates for holoprosencephaly based on expression in the developing human fetal brain, expression during the correct spatio-temporal periods of neurulation in mice embryogenesis, and/or role in developmental pathways implicated in the holoprosencephaly phenotype. In order to further identify phenotypes associated with specific genes on chromosome 14, analysis of the translocation chromosome breakpoint junctions is carried out on three patients who demonstrated phenotypes partly similar to the deletion cases analyzed, or had a breakpoint junction in the HPE8 minimal critical region. This research contributes to understanding the molecular pathogenesis of holoprosencephaly that may prove useful in the diagnosis of those affected.

Acknowledgements

I am grateful to my supervisor, Dr. Diane Cox, for her mentorship and giving me the opportunity to undertake research in her laboratory. I also express my gratitude to Diane for providing me with the freedom to scientifically explore and to grow stronger as a graduate student. I sincerely thank the members of my supervisory committee, Drs. Mike Walter, Steve Bamforth and Alan Underhill, for their suggestions on my dissertation over the years. I thank Drs. John Bell and Stephen Scherer for serving as my external examiners.

I fully thank the following people and cell bank for providing me clinical information, cytogenetic information, and blood/cell line specimens from patients and/or parents for mapping:- Drs. Maximillan Muenke (National Institute of Health, USA), Elaine Zackai (Pennsylvania School of Medicine,USA),Walter Muir (University of Edinburgh, UK), Chih-Ping Chen (Mackay Memorial Hospital, China), James Lupski (Baylor College of Medicine, USA), Paola Grammatico (University of Rome "La Sapienza,", Italy), Oliver Quarrell (Sheffield Children's Hospital, UK), Simone Schuffenhauer (Ludwig-Maximilians University Munich, Germany), Jennifer Gerritsen (Alberta Children's Hospital, Canada), Ross McLeod (Alberta Children's Hospital, Canada), John Tolmie (University of Glasgow, UK), Koenraad Devriendt (University of Leuven, Belgium), Laskhmi Mehta (North Shore Hospital, USA), and the European Collection of Cell Cultures (UK). Without these resources, my dissertation would not have been possible.

I also thank Claire Chapman of the Oxford Chromosome Abnormality database (UK) for providing me with a list of other proximal chromosome 14q constitutional and

acquired cases, and Nancy Cracknell (The Hospital for Sick Children, Toronto) for establishing lymphoblast cell lines of selected patients and parents. I am very grateful to Patricia O'Brien with Dr. Malcolm Ferguson-Smith (Cambridge University, UK) for the isolation of normal and aberrant chromosomes of selected patients by flow cytometry. I am grateful to the Cytogenetics facility (University of Alberta Hospital, Edmonton) for providing me with karyotype information and Drs. Benjamin Picard and Walter Muir for providing me with metaphase and fiber-FISH data, on selected patients used in this dissertation.

I am thankful to Mandeep Seckon and Sean Humphray (The Sanger Center, UK) for providing me with detailed assistance on the use of the FPC software and supplemental fingerprint information on selected BAC clones. I thank Seldon Briand (Canadian Bioinformatics Resource Center, Canada) for writing Perl scripts for the FPC software and database, and letting me use the UNIX server from the Canadian Bioinformatics Resource Center.

I am very grateful to Dr. Peter Dickie (University of Alberta Transgenics Centre) for teaching me dissection techniques and also for providing me with C57BL/6J pregnant mice for harvesting embryos. I am also very grateful to my colleague, Dr. Michael Wride (Cardiff University, UK) for his guidance on whole mount *in situ* hybridization. I thank Franco Joseph and Dr. Alan Underhill for letting me use their microscope facility.

I thank several past and present members of Diane's lab for their assistance over the years. Jai Shah and Aaron Davidson have assisted me with sequencing and/or the annotation of genes on chromosome 14.

This work would not have been possible without the funding from several studentship agencies. I express my sincere gratitude to these agencies that have provided me the following awards over the years: CIHR doctoral research award, AHFMR doctoral studentship award, AHFMR incentive award, Burroughs Wellcome Trust Fund student bursary, Merck Genome Research student bursary, 75th Anniversary and G.H Woods Award and Walter H. Johns Graduate Fellowship.

I finally express my infinite gratitude to my mom, dad and sister for their continual support and care during my graduate years.

Table of Contents

CHAPTER ONE: Introduction 1
1.1. Historical perspective on holoprosencephaly 2
1.2. Origin of the term "holoprosencephaly" 3
1.3. Human central nervous system and craniofacial embryology 4
    1.3.1. Central nervous system 4
    1.3.2. Face 7
    1.3.3. Eye 9
1.4. Embryonic mishaps as the cause of the holoprosencephaly phenotype 9
1.5. Holoprosencephaly forebrain defect spectrum 11
1.6. Shortcomings of associating acallosal defects with the holoprosencephaly forebrain spectrum 13
1.7. Holoprosencephaly facial spectrum 14
1.8. Extracraniofacial findings among holoprosencephaly patients 15
1.9. Other diseases associated with holoprosencephaly 16
1.10. Epidemiological findings 17
1.11. Etiology of holoprosencephaly 18
    1.11.1. Evidence for teratogenic influences 18
    1.11.2. Evidence for mechanical disruptions 19
    1.11.3. Evidence for genetic basis 20
1.12. Molecular pathways and holoprosencephaly genes 21
    1.12.1. Definition of holoprosencephaly loci 21
    1.12.2. Sonic hedgehog signaling pathway 24
        1.12.2.1. Sonic Hedgehog gene (*SHH*) 26
        1.12.2.2. Patched-1 receptor gene (*PTCH*) 27
    1.12.3. Cholesterol biosynthesis 28
        1.12.3.1. Lanosterol synthetase gene (*LS*) 29
    1.12.4. Odd paired signaling pathway 29
        1.12.4.1. Zinc finger protein in cerebellum 2 gene (*ZIC2*) 30
    1.12.5. Sine oculis pathway 31
        1.12.5.1. Sine oculis 3 gene (*SIX3*) 31
    1.12.6. Transforming growth factor beta signaling pathway 32
        1.12.6.1. TG interacting factor gene (*TGIF*) 32
        1.12.6.2. Teratocarcinoma derived growth factor 1 gene (*TDGF1*) 33
    1.12.7. WNT signaling pathway 34
        1.12.7.1. Dickkopf gene (*DKK1*) 34
    1.12.8. Retinoic acid signaling pathway 35
    1.12.9. Mutations in HPE patients are sometimes found in more than one gene 35
    1.12.10. Proposed candidate genes for holoprosencephaly 36
1.13. Mouse as a model organism to study holoprosencephaly 36
1.14. Preamble to the studies of this thesis and presentation of the general hypothesis 37

CHAPTER TWO: Clinical description and mapping of patients with proximal chromosome 14q rearrangements 42
2.1. Introduction 43
  2.1.1. The association of genetic syndromes with proximal chromosome 14q 44
  2.1.2. The association of psychiatric illnesses with proximal chromosome 14q 44
2.2 Materials and methods 45
  2.2.1. Search for proximal 14q chromosome aberrations 45
  2.2.2. Flow sorting aberrant chromosomes from lymphoblast cell lines 45
  2.2.3. Mapping aberrant breakpoints 46
  2.2.4. Determining the parent of origin for deletions 47
  2.2.5. Uniparental disomy analysis 47
  2.2.6. Microdeletion analysis 48
2.3. Clinical description of patients used in mapping 48
2.4. Results 55
  2.4.1. Mapping of aberrant chromosome breakpoints 55
  2.4.2. Defining the HPE8 minimal critical region 59
  2.4.3. Sex, ethnicity, and parent of origin studies 62
  2.4.4. Analysis of family with a Robertsonian t(14q;22q) chromosome 64
    2.4.4.1. Karyotype analysis on case 16 and selected family members 64
    2.4.4.2. Search for microdeletions 65
    2.4.4.3. Search for uniparental disomy 65
2.5. Discussion 65
  2.5.1. Mapping of chromosome rearrangements 65
  2.5.2. Parent of origin studies 68
  2.5.3. Properties of the HPE8 minimal critical region 69
  2.5.4. Family with Robertsonian t(14q;22q) chromosome 71
  2.5.5. Phenotype correlations among selected deletion cases and tentative candidate genes 72
  2.6. Summary of findings 75

CHAPTER THREE: Construction of physical and transcript maps of the holoprosencephaly minimal critical region 76
3.1. Introduction 77
  3.1.1. The physical and transcript mapping of other holoprosencephaly loci 77
  3.1.2. Physical contigs at 14q13 79
  3.1.3. Transcript maps at 14q13 80
3.2. Materials and methods 80
  3.2.1. Construction and mapping on YAC physical contig 80
  3.2.2. Construction of *in silico* BAC contig using the FPC software 81
  3.2.3. Production of BAC library filters for screening 82
  3.2.4. Production of probes and screening of BAC library filters 83
  3.2.5. Annotation of genomic sequence 84

3.2.6. RACE analysis 85
3.2.7. RT-PCR analyses on human fetal brain Poly A RNA 87
3.3. Results 88
3.3.1. Construction of a YAC contig 88
3.3.2. Construction of a BAC contig 88
3.3.3. Construction of a transcript map 90
3.3.4. Examination of the D14S1014 to AFM205XG5 interval for potential holoprosencephaly candidates 97
3.4. Discussion 99
3.4.1. Construction of physical contigs at the holoprosencephaly minimal critical region 99
3.4.2. Construction of a transcript map of the holoprosencephaly minimal critical region 100
3.5. Summary of findings 102

CHAPTER FOUR: Characterization and functional analyses of holoprosencephaly candidate genes 103
4.1. Introduction 104
4.1.1. Previous proposed candidate genes for holoprosencephaly on chromosome 14 104
4.2. Materials and methods 107
4.2.1. Screening of λ TriplEx cDNA library 107
4.2.2. Genomic structure determination 109
4.2.3. Predicted functions by bioinformatics 110
4.2.4. Construction of enhanced green fluorescent protein plasmids 110
4.2.5. RT-PCR analysis on cell lines used for transfections 112
4.2.6. Transfections 113
4.2.7. *In silico* analysis of gene expression 114
4.2.8. Human northern analysis 114
4.2.9. Extraction of Total RNA from mouse embryos 115
4.2.10. RT-PCR analysis of mouse embryonic Total RNA 116
4.2.11. Construction of probes for *in situ* hybridization 117
4.2.12. Test of sense and antisense probe expression 118
4.2.13. Production and spot testing of Riboprobes 119
4.2.14. Harvesting of mouse embryos for *in situ* hybridization 120
4.2.15. Whole mount *in situ* hybridization on mouse embryos 121
4.3. Results 123
4.3.1. Isolation of cDNA sequences 123
4.3.2. Determination of the genomic structures 125
4.3.3. Predicted properties of the proteins 129
4.3.4. Protein localization studies 131
4.3.5. Human expression profile studies 133
4.3.6. Mouse embryonic expression profile studies by RT-PCR 135
4.3.7. Whole mount *in situ* hybridization studies 139
4.4. Discussion 139
4.4.1. Characteristics of the 14q13 holoprosencephaly candidate genes 139

        4.4.1.1. *C14ORF19* (chromosome 14 open reading frame 19) 141
        4.4.1.2. *C14ORF10* (chromosome 14 open reading frame 10) 142
        4.4.1.3. *SNX6* (nexin sorting protein 6) 142
        4.4.1.4. *C14ORF11* (chromosome 14 open reading frame 11) 145
        4.4.1.5. *NPAS3* (neuronal PAS3) 146
4.5. Summary of findings 149

CHAPTER FIVE: Analysis of candidate genes for rearrangements in selected patients with derivative 14 translocation chromosomes 150
5.1. Introduction 151
5.2. Section A -- analysis of t(4;14) case 10 152
    5.3.1. The *MIPOL1* gene 152
    5.3.2. The *PITX2* gene 152
    5.4. Methods and materials 155
        5.4.1. Analysis of genomic sequence for genes at translocation breakpoint junctions 155
        5.4.2. High resolution breakpoint mapping 155
        5.4.3. Analysis of candidate genes for micro-rearrangement using flow sorted chromosomes 155
        5.4.4. Physical mapping of the *PITX2* gene 156
        5.4.5. Analysis of the *PITX2* gene for micro-rearrangement by densitometric analysis 156
    5.5. Results 157
        5.5.1. Brief clinical description 157
        5.5.2. Fine breakpoint junction mapping 157
        5.5.3. Analysis of candidate genes for micro-rearrangements 158
    5.6. Discussion 163
5.7. Section B -- analysis of familial t(9;14) cases 11 and 12 168
    5.8. Methods and materials 168
        5.8.1. Analysis of genomic sequence for genes at translocation breakpoint junctions 168
        5.8.2. Fine breakpoint mapping 169
    5.9. Results 170
        5.9.1. Brief clinical description 170
        5.9.2. Fine breakpoint mapping on chromosome 14 170
        5.9.3. Fine breakpoint mapping on chromosome 9 172
    5.10. Discussion 175
5.11. Summary of findings 178

CHAPTER SIX: General discussion 179
6.1. The hypothesis and summary of conclusions 180
6.2. Parallels between the chromosome 14q13 and 13q32 HPE loci 188
6.3. A chromosome 14q13 deletion syndrome is unlikely the cause of the phenotype 189
6.4. Models of how HPE8 candidate genes cause a holoprosencephaly phenotype 190
6.5. Significance of this dissertation 193

6.6. Future directions 194

REFERENCES 198

APPENDICES 231

Appendix 1 Contact information on patient specimens 231
Appendix 2 Markers used for mapping chromosome breakpoints in patients 232
Appendix 3 Markers used for uniparental disomy analysis of case 16 235
Appendix 4 Markers used for mapping on CEPH YAC clones at HPE8 locus 236
Appendix 5 BAC clones used for construction of physical map 236
Appendix 6 PCR primer pairs used to make probes for screening BAC filters 237
Appendix 7 PCR primers used for obtaining cDNA by RACE 238
Appendix 8 PCR primers used for RT-PCR analyses 239
Appendix 9 PCR primers used to make probes for screening λ TriplEX cDNA Library 240
Appendix 10 PCR primers used to make EGFP-N1 constructs 240
Appendix 11 Assembled cDNA sequences for HPE candidate genes 241
Appendix 12 Protein sequences of the holoprosencephaly candidate genes 243
Appendix 13 PCR primers used in chromosome 14 fine breakpoint mapping of case 10 245
Appendix 14 PCR primers used for chromosome 4 analysis of case 10 246
Appendix 15 PCR primers of candidate genes tested for rearrangement in case 10 247
Appendix 16 PCR primers used for fine breakpoint mapping of cases 11 and 12 248

List of Tables

Table 1-1: Proposed genome wide holoprosencephaly loci and genes 22
Table 2-1: Patient mapping information, listing markers in 14q1-q2 56,57
Table 2-2: Other proximal human chromosome 14q deletion and translocation Cases 61
Table 2-3: Parent of origin analyses of cases mapped 63
Table 2-4: Microdeletion and uniparental disomy analyses of case 16 67
Table 3-1: Estimated sizes of BAC clones from the HPE8 physical contig tiling path 93
Table 3-2: Summary of genes found by annotating the 2 Mb of genomic sequence from D14S49 to D14S1014 (HPE8 candidate region) 95
Table 3-3: Expression analyses of genes mapping between D14S1014 and AFM205XG5 98
Table 4-1: Genomic structure characteristics of four candidate HPE genes 127
Table 4-2: Predicted functional properties of four HPE candidate proteins based on bioinformatic analysis 130
Table 4-3: *In silico* expression profile of three candidate HPE genes 136

List of Figures

Figure 1-1: Selected illustrations on craniofacial development 5
Figure 1-2: Possible interactions among genetic pathways associated with holoprosencephaly 25
Figure 1-3: Alignment of cytogenetic intervals of chromosome 14 aberration cases associated with the holoprosencephaly spectrum 39
Figure 2-1: Examples of microsatellite typing and flow sorted chromosome analysis on selected patients 58
Figure 2-2: Definition of the HPE8 locus 60
Figure 2-3: Family with Robertsonian translocation (14q22q) chromosome 66
Figure 3-1: Three Mb YAC contig spanning the HPE8 minimal critical region 89
Figure 3-2: Construction of an *in silico* BAC contig of the HPE8 minimal critical region using the FPC software 91
Figure 3-3: Two Mb BAC contig of the HPE8 minimal critical region 92
Figure 3-4: Transcript map of 34 genes annotated in the HPE8 locus 96
Figure 4-1: Summary of information on the HPE8 locus 124
Figure 4-2: Genomic structure of HPE candidate genes 126
Figure 4-3: Expression studies of human HPE candidate genes in cell lines by RT-PCR analysis 132
Figure 4-4: Cellular localization studies of HPE candidate proteins 134
Figure 4-5: Human northern analysis of HPE candidate genes 137
Figure 4-6: Expression studies by RT-PCR on the HPE8 candidate murine holoprosencephaly orthologues 138
Figure 4-7: Expression studies by whole mount *in situ* hybridization on selected day 11 C57BL/6J murine embryos 140
Figure 5-1: Breakpoint junction mapping of case 10 using flow sorted chromosomes 159
Figure 5-2: Analysis of the *MIPOL1* gene in case 10 for rearrangements 160
Figure 5-3: Flow sorted chromosome analysis with markers mapping within a panel of candidate genes flanking the translocation breakpoint junctions on chromosomes 14q13 and 4q25 in case 10 162
Figure 5-4: Analysis of the *PITX2* gene for rearrangement in case 10 164
Figure 5-5: Breakpoint junction analysis of chromosomes 14 and 9 of cases 11 and 12 171
Figure 5-6: Search for rearrangement in the *NPAS3* gene in the t(9;14) family 173
Figure 5-7: Identification of the rearrangement in the *NPAS3* gene in the t(9;14) family 174

List of Symbols, Nomenclature, or Abbreviations

μM	micromole
μL	microlitre
ACC	agenesis of the corpus callosum
AD-003	AD-003 protein gene
AHC	yeast complete medium
ANK2	ankyrin 2 gene
ARS1	Axenfeld Reiger syndrome type 1
BAC	bacterial artificial chromosome
BAZ1A	bromodomain adjacent to zinc finger domain 1A gene
BF1 (FOXG1)	brain factor 1 gene
bHLH	basic helix loop helix
BLAST	basic local alignment search tool
BMP4	bone morphogenetic protein 4 gene
C14ORF10	chromosome 14 open reading frame 10 gene
C14ORF11	chromosome 14 open reading frame 11 gene
C14ORF19	chromosome 14 open reading frame 19 gene
cDNA	complementary DNA
CEPH	Centre d'Etude du Polymorphisme Humain
CFL2	cofilin 2 gene
CNS	central nervous system
COCH	cochlin gene
CRAT	carnitine acetyltransferase gene
CT	computed tomography
Dach	dachshund gene
ddH$_2$0	double distilled water
DKK1	dickkopf gene
DNA	deoxyribonucleic acid
DOP-PCR	degenerate oligonucleotide primed polymerase chain reaction
dATP	deoxyadenosine tri phosphate
dCTP	deoxycytidine tri phosphate
dGTP	deoxyguanosine tri phosphate
dNTP	deoxynucleoside tri phosphate
dpc	days post coitus
dTTP	deoxythymidine tri phosphate
E	embryonic stage
EEG	electroencephalocardiogram
EGF	epidermal growth factor gene
EGF-CFC	epidermal growth factor-like; Cripto, Frl-1, Cryptic domain
EGLN3	egl nine homolog 3 gene
EGFP	enhanced green fluorescent protein
En	engrailed gene
EST	expressed sequence tag
Ey	eyeless gene
FISH	fluorescent in situ hybridization
FPC	fingerprinted contig

GAPDH	glyceraldehyde-3-phosphate dehydrogenase gene
GB4	gene bridge 4
GI	genbank identifier number
GLI3	glioma 3 gene
HNF3A (FOXA1)	hepatocyte nuclear factor 3 A gene
HPE	holoprosencephaly
HPE8	holoprosencephaly locus 8
IRD	infrared dye
JNK	c-Jun N terminal kinase
kDa	kilodalton
LB	Luria-Bertani broth
Ls	lanosterol synthetase gene
LSFR2	linked to Surfeit genes in Fugu rubripes 2 gene
Mb	megabase
Mbq	megabecquerel
mCi	milli curie
MIH	middle interhemispheric variant
MIPOL1	mirror image polydactyly 1 gene
mM	micromole
MRI	magnetic resonance imaging
NCBI	national centre for biotechnology information
NFKBIA	nuclear factor kappa B inhibitor of alpha gene
NKX2.1 (TTF1)	thyroid transcription factor 1 gene
NKX2.9	NK homeobox 2.9 gene
NMDAR1	N-methyl D-aspartate receptor 1
NPAS3	neuronal PAS3 gene
OD	optical density
OFC	occipto frontal circumference
Opa	odd paired gene family
ORF	open reading frame
PAC	P1 artificial chromosome
PAS	period, aryl hydrocarbon receptor, single minded
PAX9	paired homeobox 9 gene
PCR	polymerase chain reaction
PFU	plaque forming unit
PITX2	paired like homeodomain transcription factor 2 gene
PPP2R4	protein phosphatase 2A, regulatory subunit B gene
PRO0971	hypothetical protein PRO0971 gene
PSMA6	proteosome macropain subunit alpha 6 gene
PTCH	patched-1 receptor gene
RACE	rapid amplification of cDNA ends
RNA	ribonucleic acid
RPCI	roswell park cancer institue
RT-PCR	reverse transcriptase polymerase chain reaction
RXR	retinoid receptor response element
SHH	sonic hedgehog gene

Sim	single minded gene
SIM2	single minded homolog 2 gene
SIX	sine oculus
SIX3	sine oculus 3 gene
SLC25A21	solute carrier family 25 member 21 gene
SMAD	mothers against decapentaplegic
SMAD1	mothers against decapentaplegic homolog 1
SMAD2	mothers against decapentaplegic homolog 2
SNX6	nexin sorting protein 6 gene
So	sine oculus
SRP54	signal recognition particle 54 gene
STS	sequenced tagged site
Wg	wingless gene
T2BP	hypothetical protein MGC20791 gene
TCR	T-cell receptor gene cluster
TDGF1	tetratocarcinoma derived growth factor 1 gene
TGFß	transforming growth factor beta
TGIF	TG interacting factor gene
UPD	uniparental disomy
UTP	uridine triphosphate
UTR	untranslated region
UV	ultraviolet
WIGR	whitehead institute for genome research
YAC	yeast artificial chromosome
YPD	yeast extract/peptone/dextrose medium
ZIC	zinc finger protein in cerebellum gene family
ZIC2	zinc finger protein in cerebellum 2 gene

CHAPTER ONE

INTRODUCTION

1.1. HISTORICAL PERSPECTIVE ON HOLOPROSENCEPHALY

Over the past centuries, accounts of holoprosencephaly were closely affiliated not only with the field of teratology, but also with superstition, and fables. Such accounts were reported in the works of poets, artists and even philosophers. Many myths of this disease were initially focused on one-eyed people or Cyclops, particularly in Greek mythology where Polyphemus was the giant, ugly, man-eating cyclopic shepherd from Sicily. Other cultures, such as the Sri Lankan's, viewed the Cyclop as a supreme spiritual entity to extirpate evil spirits (Cohen and Sulik, 1992), while the Babylonian tablets in the British Museum described one-eyed beast like people who lived in the Altai mountains of northeast Scythia and stole gold from the griffins (Warkany, 1971). The modern western culture has perceived the cyclopic phenotype as a figure for artistic expression, most notably in the painting "The Difficult Crossing", by Rene Magritte (Cohen and Sulik, 1992).

Unfortunately, people who were born with the holoprosencephaly spectrum of phenotypes, particularly cyclopia, were initially thought to be an omen from God's anger or the demon. As a result, people with such dysmorphisms were punished or executed. It was not until the early period of data collection, data analyses and the advent of the field of teratology by people such as Etienne Geoffroy Saint-Hilaire and Willem Vrolik that scientific thought introduced rationality to the holoprosencephaly phenotype (Cohen and Sulik, 1992). The old scheme of thought by the early teratologists proposed that the holoprosencephaly phenotype was due to the effect of atypical physical conditions upon the developing fetus. In this respect, the parents were to be blamed since their lifestyles could alter the outcome for the developing child. Nonetheless, the refinement of the cause of the holoprosencephaly phenotype and the classification of the phenotype into groups or classes assisted in the evolution of our current perspectives on this most interesting disease.

1.2. ORIGIN OF THE TERM "HOLOPROSENCEPHALY"

The earliest classification system for the holoprosencephaly spectrum was from the works of Isidore Geoffroy Saint-Hilaire (Cohen and Sulik, 1992). His scientific terms were principally used to describe the varying degrees of ocular fusion in addition to defects of the proboscis, mandible and stomodeum. He coined the terms ethmocephalus and cebocephalus to describe two closely set orbital fossae; and rhinocephalus, cyclocephalus and stomocephalus to describe a single orbital fossae. Since the olfactory tracts and bulbs were affected among these subjects, the term arhinencephaly was later suggested by Kundrat, to include these affected structures with the central nervous system anomalies (Cohen and Sulik, 1992). The term arhinencephaly was nonetheless replaced with holotelencephaly by Yakovlev, to describe the failure of the prosencephalon to form the cerebral hemispheres with separate lateral ventricles (Cohen and Sulik, 1992). It was not until 1963 that DeMyer and Zeman initially proposed the term holoprosencephaly (HPE) to collectively include the central nervous system spectrum of anomalies. In this manner, holoprosencephaly referred to the impaired patterning or "cleavage of the prosencephalon sagitally into cerebral hemispheres, transversely into telencephalon and diencephalon, and horizontally into olfactory and optic bulbs" (DeMyer and Zeman, 1963).

1.3. HUMAN CENTRAL NERVOUS SYSTEM AND CRANIOFACIAL EMBRYOLOGY

A brief description of the embryology of human central nervous system and craniofacial embryology is outlined below. Many books and articles have been written on this subject and should be consulted for further details (Larsen, 1993; Rana, 1998; Fix and Dudek, 1995; Kahn, 2000; Elias et al., 1991; Carstens, 2002). A description of the embryological development of the neurocranium and craniofacial muscle was omitted in this section since the main focus of the holoprosencephaly phenotype is the facial and central nervous system anomalies.

1.3.1. Central nervous system

During the post gastrulation period at the third week of gestation, the primitive streak forms within the trilaminar shaped embryo. Not all of the embryo is in a trilaminar shape. In fact, the cranial midline has a bilayer of cells. The cells subjacent to the median ectoderm at the rostral part of this bilayer produce the prechordal plate (Figure 1-1a). The position of the prechordal plate dictates where the primitive oral cavity and stomodeum develop.

As neurulation proceeds, the primitive streak patterns into the neural plate, which curves into the neural grooves up to the fourth week of gestation (Figure 1-1b). Cells at the junction between the surface ectoderm and neural plate become neural crest cells or the fourth germ layer. Neural crest cells have the potential to differentiate into pigment cells, nerve cells (sensory and motor), glial cells, and skeletal and connective tissue of the head. The multiplicity of differentiating fates originating from neural crest cells is postulated as a result of spatial programming defined within segmented regions of the

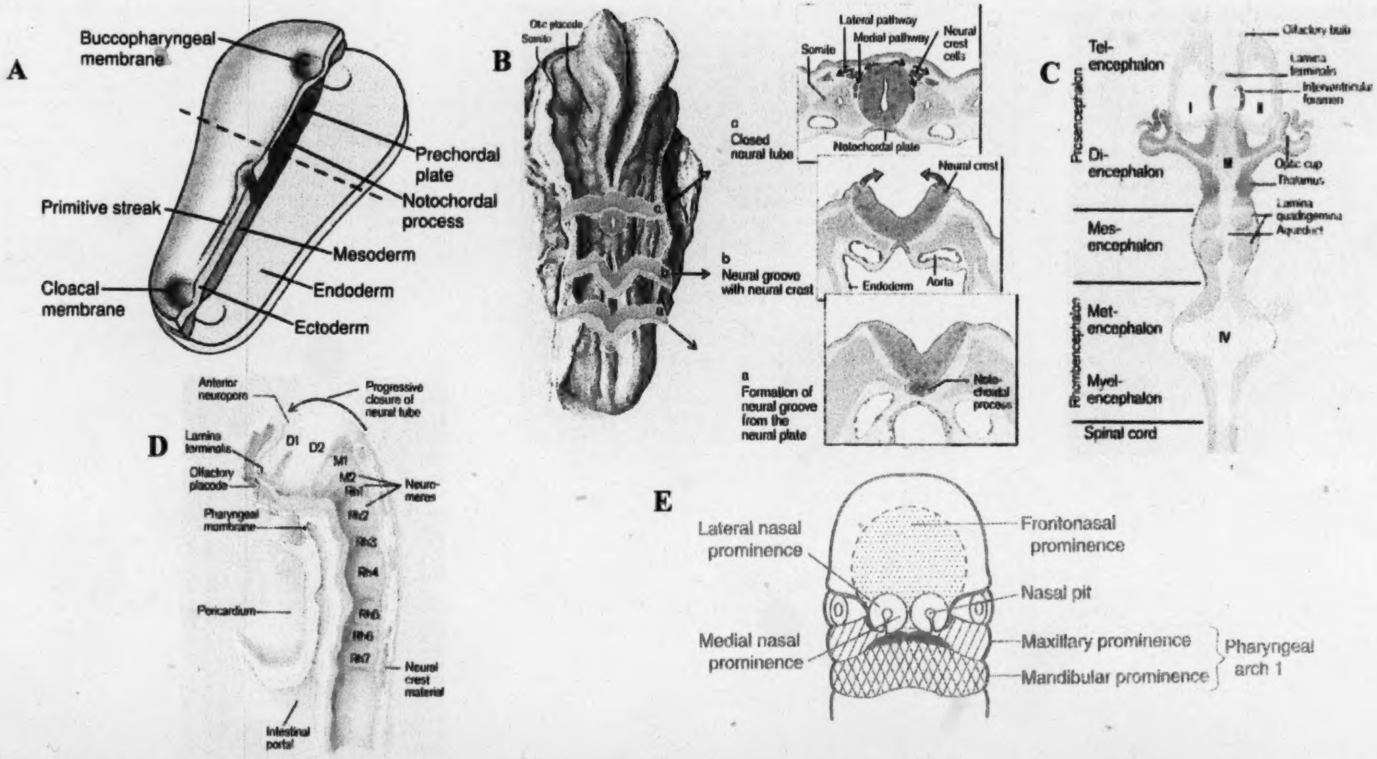

A
Buccopharyngeal membrane
Prechordal plate
Notochordal process
Primitive streak
Mesoderm
Endoderm
Ectoderm
Cloacal membrane
B
Otic placode
Somite
a Closed neural tube
b Neural groove with neural crest
a Formation of neural groove from the neural plate
Lateral pathway
Medial pathway
Neural crest cells
Notochordal plate
Neural crest
Endoderm
Aorta
Notochordal process
C
Tel-encephalon
Di-encephalon
Prosencephalon
Mes-encephalon
Met-encephalon
Myel-encephalon
Rhombencephalon
Spinal cord
Olfactory bulb
Lamina terminalis
Interventricular foramen
Optic cup
Thalamus
Lamina quadrigemina
Aqueduct
IV
D
Anterior neuropore
Progressive closure of neural tube
Lamina terminalis
Olfactory placode
Pharyngeal membrane
Pericardium
Intestinal portal
D1
D2
M1
M2
Rh1
Rh2
Rh3
Rh4
Rh5
Rh6
Rh7
Neuromeres
Neural crest material
E
Lateral nasal prominence
Medial nasal prominence
Frontonasal prominence
Nasal pit
Maxillary prominence
Mandibular prominence
Pharyngeal arch 1

Figure 1-1: Selected illustrations of craniofacial development. A) trilaminar shaped embryo during post gastrulation. B) neural folding during neurulation. C) segmentation of the brain vesicles. D) locations of neuromeres. D1-2= prosomeres (third prosomere not shown), M1-2 = mesomeres, R1-7 = rhombomeres (all 12 rhombomeres not shown). E) primitive facial structures. Note, Figure A is modified from Figure 3-5 of "Human embryology" by Larsen with permission from Churchill Livingstone. Figures B-D are modified from plates 18.1, 18.3 and 18.4 of "Human embryology made easy", by Rana with permission from Harwood academic publishers. Figure E is modified from Figure 12-4 of "Embryology", by Fix and Dudek with permission from Williams and Wilkins.

developing central nervous system under the influence of embryonic fields. In this manner, embryonic fields are defined as tissue blocks that establish a program of the developing nervous system (Carstens, 2002). Within these fields extensive formation, migration, coalescence and interaction occur among them (Carstens, 2002). By the fifth week of gestation, the neural grooves fold inward and commence to fuse and form the neural tube at the junction that later becomes the cervical cord. During this period, the neural tube at the anterior neuropore closes first, followed by the closure of the posterior neuropore about two to three days later. About two thirds of the neural tube develops into the central nervous system, and the remaining third forms the spinal cord. The ventricular system and central canal develop from the lumen of the neural tube, while the further thickening and folding of the neural tube form the rest of the brain and spinal cord.

During the fifth week of gestation, the closure of the rostral neuropore is the prerequisite for the differentiation of the primary vesicles, namely prosencephalon (primitive forebrain), mesencephalon (primitive midbrain) and rhombencephalon (primitive hindbrain) (Figure 1-1c). A second closure of the neural tube occurs in a rostro-caudal manner in the prosencephalon. At this stage, the central nervous system becomes discrete segmental units known as neuromeres (Figure 1-1d). Each neuromere acts as a genetic barcode to define the position and fate of the neuromere. At the beginning of the sixth week of gestation, the three primary vesicles undergo further segmentation. The prosencephalon has three prosomeres. Two prosomeres of the diencephalon later differentiate into the thalamus, hypothalamus, globus pallidus and third ventricle, while a single prosomere of the telencephalon later differentiates into the cerebral hemispheres, caudate nucleus, putamen and lateral ventricles. The mesencephalon has two mesomeres that develop into the midbrain and aqueduct of Sylvius. The rhombencephalon has twelve rhombomeres that divide into the metencephalon which later develops into the pons, cerebellar hemispheres, vermis, part of the fourth ventricle; and the myelencephalon which later produces the medulla and part of the fourth ventricle. By the end of the sixth week of gestation, the architecture of the brain is established.

Proceeding from the seventh week of gestation, the brain undergoes further growth and maturation. During this period, the gyri and sulci are produced, and there is extensive neuronal migration. Between the 9th and 10th week of gestation, radial glial fibres extend from the germinal matrix thereby lining the ventricles to the brain surface. Neurons migrate in an orderly fashion along the radial fibres to develop a six layered cerebral cortex. During the 20th week of gestation, the Sylvian or lateral sulcus forms. Furthermore, from 23 to 25 weeks of gestation, an increase in the folding of the brain can be noted with the formation of the rolandic, interparietal and superior temporal fissures. From 32 to 33 weeks of gestation, the secondary and tertiary sulci develop.

1.3.2. Face

The majority of the face develops essentially between the third and eighth week of gestation. Seven somites form in a cranial to caudal sequence from the developing prosencephalon to the occipital region, each one associated with a specific portion of the cranial neural plate. The somitomeric mesoderm produces muscle cells associated with the visceral arches and the upper portion of the skull. The tissue surrounding the prosencephalon becomes the frontonasal prominence. By the fifth week of gestation, localized thickening of the ectoderm on the frontolateral aspects of the frontonasal prominence surround the nasal prominences (Figure 1-1e). The nasal prominences develop into the nasal placodes in which the deepening of the medial and lateral nasal prominences produce the nasal pits. The lower portion of the medial nasal prominence merges to develop the philtrum, the upper four incisors and primary palate. On each side of the developing face, fusion of the medial nasal prominence with the lateral nasal prominence and maxillary prominence of the first pharyngeal arch completes the development of the upper lip. The two mandibular prominences form in a medioventral manner to eventually develop the lower lip and lower jaw. Furthermore, the fusion of the mandibular and maxillary prominences demarcates the opening of the stomodeum.

The pharyngeal and branchial arches develop in a cranial to caudal sequence on the ventrolateral aspect of the developing face and neck. The pharyngeal arches initially serve as the aortic arch arteries but later, each arch is associated with a specific cranial nerve, muscular or skeletal derivative. The muscular components are mesodermally derived, while the connective tissue components are derived from the neural crest cells. The maxillary and mandibular prominences are part of the first pharyngeal arch (Figure 1-1e). The maxillary prominences also produce the majority of the secondary palatal shelves by the 6th to 7th week of gestation. By the 9th week of gestation, union of the primary and secondary palatal shelves, and the nasal septum occurs. Fusions occur in an anterior to posterior direction thus separating the nasal cavity from the oral cavity. Initially, the mandibular prominences consist of mesodermal cells; however neural crest cells from the lower midbrain and upper hindbrain join the mesodermal cells to populate these prominences. The lower jaw and portion of the tongue are major derivatives of the mandibular prominences. Tissues from the mandibular prominences and hyoid arches produce the external ears and bones of the middle ear. Breakdown of the buccopharyngeal membrane results in continuity between the stomodeum and pharynx. The epithelium on the stomodeal side of the membrane is ectodermally derived while on the pharyngeal side is endodermal. At this ectodermal-endodermal boundary, the front of the palatine tonsils arises.

1.3.3. Eye

During the third week of gestation, two evaginations centrally placed on the lateral walls of the future diencephalon form the optic sulci (Figure 1-1c). In the fourth week of gestation, deepening of the optic sulci produces the optic primordia which enlarge to develop the optic vesicles as the anterior neural folds close. Interaction between the optic vesicle and surface ectoderm produces the lens placode. Invagination of the optic vesicle to develop the bilayered optic cup is accompanied by invagination of the lens placode to produce the lens vesicle. The cornea develops following the separation of the lens vesicle from the surface ectoderm. The ectoderm forms the epithelium of the eye while the remaining parts of the eye are of neural crest origin.

1.4. EMBRYONIC MISHAPS AS THE CAUSE OF THE HOLOPROSENCEPHALY PHENOTYPE

Pathological studies of the holoprosencephaly phenotype have been a major challenge since the majority of embryos die *in utero*. Cranial imaging methods by ultrasound, computed tomography (CT) and magnetic resonance imaging (MRI) have assisted to some degree in deciphering the timing, place and type of defect. However, details of defective internal structures are limited. Several groups have been successful in the prenatal diagnosis of a holoprosencephalic fetus as early as 10 weeks (Wong et al., 1999, Turner et al., 1999) to 23 weeks of gestation (Hsu et al., 2001; Bronshtein and Wiener, 1991) using transvaginal ultrasound, three dimensional MRI or helical CT. The best methods to understand the cause of the holoprosencephaly phenotype thus far have been three dimensional reconstructed images from serial histological sections or cranial scans (Cannistra et al., 2001; Shiota et al., 1993). Using these methods, several reports have outlined in much detail the structural defects apparent in holoprosencephaly concepti (Cannistra et al., 2001; Shiota et al., 1993; Kjaer et al., 2002).

Isidore Geoffroy Saint-Hilaire initially proposed that the pathogenesis of the holoprosencephaly phenotype was due to a lack of the development of the nasal apparatus and varying fusions of the orbital fossae (Cohen and Sulik, 1992). Further refinements in this scheme of thought by Dareste led to the proposal that the phenotype is due to the premature closure of the anterior neuropore due to a deficiency of anterior cerebral precursor cells and defects in mesenchymal development (Cohen and Sulik, 1992). Such a phenotype may be due to many factors such as mechanical disruptions whereby there is increased amniotic pressure against the developing prosencephalon. Modern schemes of thought propose that insults up to the seventh week and as early as the third week of gestation contribute to midline defects particularly seen in holoprosencephaly and agenesis of the corpus callosum. The pathogenesis of the holoprosencephaly phenotype is therefore a result of the disturbance of the mesencephalic neural crest, insufficient formation of structures derived from the prechordal plate, aberrant distribution of the cephalic mesenchyme, and/or defective interaction between the cephalic tip of the notochordal plate, neuroectoderm of the brain or oral plate (Webster et al., 1988; Cohen et al., 1971; Sulik and Johnston, 1982; Muller and O'Rahilly, 1989). With these proposed mechanisms, agenesis of the ethmoid bone, which develops from the prechordal plate (at week six of gestation), can cause the nasal ectoderm to not contact the ethmoidal plate, resulting in the majority of midface hypoplasia observed. These include a proboscis, fusion to varying degrees of the two optic vesicles in cyclopia, intercanthal spacing anomalies, malformed upper lip and philtrum, and an absent nasal septum (Siebert, 1981; Cannistra et al., 2001). Clefting of the palate and lip is thought to be a secondary consequence of an absent ethmoid bone (Cannistra et al., 2001). Aberrant migration of the palantine shelves from the vertical to horizontal plane and fusion at the inferior border of the perpendicular plate of the developing ethmoid bone in the premaxillary region is one of many proposed mechanisms that result in clefting (Siebert, 1981). The lack of craniofacial embryonic midline fields also contributes to single central maxillary incisor, and agenesis or hypoplasia of the associated central nervous system midline structures.

1.5. HOLOPROSENCEPHALY FOREBRAIN DEFECT SPECTRUM

The forebrain malformations are graded into alobar (54-64%), semilobar (18-24%) and lobar (9-12%) types according to the extent of severity (Olsen et al., 1997). Forebrain malformations classified as alobar HPE are further graded into pancake, cup and ball subtypes depending on the extent the dorsal lip of the holotelencephalon rolls to cover the membranous ventricular roof. The alobar HPE forebrain can be described as a small monoventricular cerebrum with an absent interhemispheric fissure and falx cerebri (Cohen, 1989a,b; Peebles 1998). Moreover, the holoprosencephalon may occupy varying positions along the anterior-posterior axis of the skull (Cohen, 1989a,b). A few commissural fibers may cross the midline, however the thalami and corpora striata are still undivided. Other anomalies include absent Sylvian fissures, presence of dorsal cyst, dysplastic basal ganglia and thalamus, hypoplastic or absent corticospinal tracts, agenesis of the corpus callosum, and absent olfactory tracts and bulbs (Cohen, 1989a,b).

Affected subjects with semilobar HPE have rudimentary cerebral lobes and an incomplete interhemispheric fissure located at the posterior (Cohen, 1989a,b). The corpora striata is continuous across the midline and the corpus callosum is absent or underdeveloped. Again, a few commissural fibers may cross the midline. Agenesis or hypoplasia of the olfactory tracts and bulbs is common. Both alobar and semilobar HPE subjects have a reduced brain volume.

Subjects with lobar HPE have well formed cerebral lobes with a distinct interhemispheric fissure, however, there is partial fusion in the parietal and frontal lobes (Peebles, 1998). The cingulate gyrus may frequently cross the midline (Cohen and Sulik, 1992). The midline cleavage of the thalami and corpora straita may be incomplete and the corpus callosum and cavum septum pellucidum may be absent (Peebles, 1998). The corpus callosum, olfactory tracts and bulbs may be absent or hypoplastic and there is variable enlargement of the lateral ventricles. The forebrain is of normal size. Since the structural malformations of a lobar HPE forebrain are subtle, it is very difficult to prenatally diagnose a fetus with lobar HPE by transvaginal ultrasound examination (Peebles, 1998).

The middle interhemispheric variant (MIH), also known as syntelencephaly, refers to the failure of separation of the posterior frontal and parietal lobes. The incidence is yet to be calculated since this is only a recently defined category of the mild holoprosencephaly forebrain spectrum of anomalies. In MIH, the midline of the basal forebrain, anterior frontal lobes and occipital lobes are properly defined. Patients demonstrating MIH show microcephaly, normal or mild midface anomalies and central nervous system anomalies including more than one of the following: partial or complete agenesis of the corpus callosum, heterotopic gray matter or cortical dysplasia, abnormal sulci, absent septum pellucidum, underdeveloped choroid plexus, underdeveloped hippocampus, incompletely separated caudate nuclei, hypoplastic or absent olfactory bulbs, underdeveloped pituitary gland, dysplastic, partial or completely fused thalami and abnormal anterior arterial circulation (Marcorelles et al., 2002; Simon et al., 2002; Barkovich and Quint, 1993). In support of MIH as part of the holoprosencephaly spectrum, many cases ascertained as lobar HPE were rediagnosed as MIH (Corona-Rivera et al., 2000; Imaizumi et al., 1998). Furthermore, a patient with MIH who was born to a diabetic mother, also had a single central maxillary incisor and an absent anterior maxillary frenulum (Robin et al., 1996), supporting MIH as part of the holoprosencephaly spectrum.

1.6. SHORTCOMINGS OF ASSOCIATING ACALLOSAL DEFECTS WITH THE HOLOPROSENCEPHALY FOREBRAIN SPECTRUM

Agenesis of the corpus callosum (ACC) is the most common brain malformation with an incidence of up to one in 20,000. The causes are extremely heterogeneous and both HPE and ACC share genetic and non-genetic etiologies (Jellinger et al., 1981). Complete and partial agenesis of the corpus callosum is part of the holoprosencephaly spectrum. As a result, it is a challenge to ascertain whether a patient with acallosal defects with or without other central nervous system and midfacial anomalies is affected with mild holoprosencephaly or isolated acallosal defects. In the case of partial agenesis of the corpus callosum, some commissural fibers are present to form the splenum of the corpus callosum. In isolated agenesis of corpus callosum, the callosal axons form at the correct site but cannot cross the midline in a posterior to dorsomedial direction in the cerebral hemispheres (Dobyns, 1996). When this occurs, the callosal axons aggregate to form the bundle of Probst that lies along the supermedial aspect of the lateral ventricles. In holoprosencephaly, the bundle of Probst is not seen. Callosal axons do form but at aberrant sites thus preventing them from following the correct migration pathway (Rubinstein et al., 1996; Sarnat and Flores-Sarnat, 2001). The interhemispheric fissure is proposed to act as a bridge for the first few callosal axons to cross, thus serving to mediate a passage for the crossing of millions of callosal axons across the midline (Rubinstein et al., 1996). For partial agenesis of the corpus callosum in HPE, a normal dorsal commissure plate can be induced and alternate routes of callosal axon migration, particularly in the posterior CNS (central nervous system) can lead to corpus callosum development (Rubinstein et al., 1996). It is also suggested that in lobar HPE cases, there is posterior displacement of the lamina terminalis and commissural plate and the rostrum of the corpus callosum is hidden in white matter, resulting in the false appearance of partial agenesis of the corpus callosum (Schaefer et al., 1991). But this idea is refuted by MRI and other autopsy studies (Rubinstein et al., 1996; Cohen and Sulik, 1992).

1.7. HOLOPROSENCEPHALY FACIAL SPECTRUM

The forebrain defects are commonly associated with facial dysmorphisms. In fact, there is a positive correlation in an estimated 80% of the cases between the severity of the facial and forebrain defects (DeMyer et al., 1964). The facial spectrum includes cyclopia (median monopthalmia, synopthalmia or anopthalmia with or without probosis formation and arhinia), ethmocephaly (hypotelorism with arhinia and probosis formation), cebocephaly (hypotelorism and absent columella) and median cleft lip (Cohen and Sulik, 1992). Other less severe facial findings may include hypotelorism, hypertelorism, lateral cleft lip, flat nose, iris coloboma, single maxillary central incisor, agenesis of the nasal septal cartilage, stenosis of the pyriform aperture, and agenesis of the philtral ridges (Hennekam et al., 1991; Tavin et al., 1994; Ming and Muenke, 1998). The presence of a single central maxillary incisor is thought to increase the risk for an HPE conceptus if this trait is manifested in one of the parents. However, it is noteworthy that a single central maxillary incisor can occur as an isolated anomaly with an estimated incidence of 1/50000, or is manifested with other non-holoprosencephaly diseases (Cohen, 2001; Becktor et al., 2001).

1.8. EXTRACRANIOFACIAL FINDINGS AMONG HOLOPROSENCEPHALY PATIENTS

Individuals affected with holoprosencephaly typically have malformed cranial sutures within the craniosynostosis spectrum. These include skulls that are microcephalic, hydrocephalic, trignocephalic, or brachycephalic (Cohen and Sulik, 1992; Marcorelles et al., 2002). Many other anomalies not associated with the head are also evident. These include defects of the extracranial skeleton (congenital dislocation of hips, short sternum, vertebral anomalies), body wall (omphaloceol, sacral dimple), upper and lower extremities (polydactyly, syndactyly, clinodactyly, camptodactyly), skeletal muscles (malformed tongue, diaphragmatic hernia), respiratory system (choanal stenosis, choanal atresia, pulmonary hypoplasia), cardiovascular system (atrial septal defect, ventricular septal defect), gastrointestinal tract (tracheoesophageal fistula, malrotation of intestine, agenesis of gall bladder, malformed pancreas), genitourinary system (hypoplastic penis, vaginal atresia, renal hypoplasia/agenesis, urethral stenosis), reticuloendothelial system (malformed thymus and/or spleen), endocrine system (malformed pituitary, thyroid gland and/or adrenal gland, cryptoorchidism) (Cohen, 1989 a,b; Cohen and Sulik, 1992). Many of the above anomalies have received very minimal attention with ascertainment of the holoprosencephaly spectrum of phenotype, and therefore warrant further studies. In general, there is tremendous clinical heterogeneity of the holoprosencephaly spectrum in affected subjects.

1.9. OTHER DISORDERS ASSOCIATED WITH HOLOPROSENCEPHALY

An estimated 18 to 25% of HPE cases can be associated with more than 60 defined monogenic syndromes (Cohen, 1989a; Cohen and Sulik, 1992; Armbruster-Moraes et al., 1999; Lapunzina et al., 2001). Other disorders such as Klinefelter syndrome (Armbruster-Moraes, et al., 1999), Genoa syndrome (Lapunzina et al., 2001), Chiari II malformation (Rollins et al., 1999), schizophrenia (Hercig et al., 1994; Roach et al., 1975), bilateral vocal cord palsy (Smilari et al., 2001), frontoethmoidal encephalocele (Elgin et al., 2001), exadactyly (Karantanas et al., 1999), split hand deformity (Abdel-Meguid and Ashour, 2001) and profound deafness (Vantrappen et al., 1999) have also been associated with holoprosencephaly. The underlying cause of these associations is unknown. However, defects in more than one gene and gene-environment interactions can likely result in holoprosencephaly associated with these disorders.

1.10. EPIDEMIOLOGICAL FINDINGS

The incidence of holoprosencephaly varies from 1 in 5,200 to 53,395 (Corsello et al., 1990; Forrester and Merz, 2000; Olsen et al., 1997). A more accurate estimate of the prevalence is 0.86 to 1.2 per 10,000 livebirths, based on studies from the Japanese, American and Spanish populations (Matsunga and Shiota, 1977; Roach et al., 1975; Urioste et al., 1988). The underlying problems with discerning the prevalence of HPE in the population are due to factors such as incomplete ascertainment, incomplete clinical records, differing type of population (newborn verses stillborn), and inclusion of only certain types of holoprosencephaly (Cohen, 1989a). The incidence is thought to be higher since about 40 in 10,000 aborted concepti are thought to be holoprosencephalic (Matsunaga and Shiota, 1977) and many cases of children or fetuses with normal or mild facial findings but an uncharacterized forebrain anomaly are not included in the incidence calculation. Interestingly, the rate of HPE has increased from 0.58 per 10,000 from 1968 to 1972, to 1.2 per 10,000 from 1988 to 1992 with an unknown cause (Cohen 1989a; Rasmussen et al., 1996). The incidence is higher in non-caucasians (Asian, Hispanic, and African population) than in caucasians (Matsunaga and Shiota, 1977; Forrester and Merz, 2000; Olsen et al., 1997; Whiteford and Tolmie, 1996). Moreover, among the races, there is extreme intra and inter familial clinical variability (Corsello et al., 1990; Whiteford and Tolmie, 1996). Epidemiologic studies have further suggested factors such as poverty and exposure to teratogens demonstrate a significant link to the genesis of holoprosencephaly (Cohen, 1989a; Croen et al., 2000).

Studies on the sex distribution of holoprosencephaly have unexpectantly found up to a 3:1 female to male ratio for alobar HPE, and a 1:1 female to male ratio for lobar HPE

(Cohen, 1989a,b; Croen et al., 1996). Since severe facial types are associated mostly with severe forebrain findings, affected females frequently have more severe facial anomalies than do males (Cohen, 1989a,b). In general, females manifest the holoprosencephaly trait more than males (Rasmussen et al., 1996; Olsen, et al., 1997). It is yet to be determined whether specific X-linked modifier genes are the cause of the sex bias. No strong evidence for maternal age effect has yet been correlated with the cause of holoprosencephaly.

Among specific craniofacial anomalies, eye malformations (micropthalmia, hypotelorism) are the most common followed by nose (flat nose, excess proboscis), ear (low set, deformed, rotated, large, small) and oral clefts in the study of a New York population (Olsen et al., 1997). Clefting is most commonly associated with eye and nose malformations (Olsen et al., 1997). There is also an inverse correlation in the incidence of facial anomalies and its severity (Croen et al., 1996). Cyclopia, ethmocephaly and cebocephaly occur less than 50% (Croen et al., 1996).

1.11. ETIOLOGY OF HOLOPROSENCEPHALY

The etiology of holoprosencephaly is heterogeneous, with teratogenic influences, mechanical embryonic insults and a genetic basis as postulated causes.

1.11.1. Evidence for teratogenic influences

Spontaneously induced holoprosencephaly phenotype in animals such as fish, frogs, mice, chick, guinea pigs, sheep, rabbits and rats have suggested an assortment of chemical and physical teratogens, including retinoic acid, ethyl alcohol, salicylate, veratrum alkaloids, cholesterol biosynthesis inhibitors, infections (rubella, cytomegalovirus, toxoplasmosis virus) or irradiation can variably contribute to the pathogenesis of holoprosencephaly (Cohen, 1989b; Cohen and Sulik, 1992). The variable expressivity of the holoprosencephaly phenotype is due to differences in the timing, duration or potency of the teratogen on embryogenesis (Cohen and Sulik, 1992). Pregnant mothers with diabetes have been observed to have a 200 fold increased risk for having a holoprosencephalic conceptus (Barr et al., 1983). Epidemiological studies previously demonstrated that alcohol, cigarette smoking, and even salicylates can increase the risk for holoprosencephaly in humans (Croen et al., 2000). The use of tetratogens has been important in comprehending the embryological cause of holoprosencephaly in animal models. Spontaneously induced holoprosencephaly suggested its genesis can commence as early as gastrulation, which is 3 to 4 weeks of gestation in humans, or 6 to 7 gestational days in mice (Webster et al., 1988). Further studies are needed to demonstrate the gene-environment interactions, and the specific timing and intensity of the insult on the penetrance and expressivity of holoprosencephaly.

1.11.2. Evidence for mechanical disruptions

Constriction of early cleavage or excision of the prechordal mesoderm from the gastrula has been shown to induce holoprosencephaly in frogs (Cohen and Sulik, 1992). Some of these early findings led to suggest that the failure of the prechordal mesoderm to interact with the overlying surface ectoderm can serve as an embryonic insult in the genesis of holoprosencephaly (Webster et al., 1988). Amniotic banding has also been associated with the holoprosencephaly phenotype, however the actual mechanism is yet to be elucidated (Hudgins et al., 1985,1986). The lack of proper vascularization by means of aberrant middle and anterior cerebral arteries, as seen among holoprosencephaly patients, is another possibility for pathogenesis of this disease. With this mechanism, nutrients to support the establishment of embryonic fields are inadequate (Carstens, 2002; Bernard et al., 2002).

1.11.3. Evidence for genetic basis

Several lines of evidence suggest a genetic basis for holoprosencephaly. Firstly, the features of the holoprosencephaly spectrum can often be part of a syndrome (DeMeyer and Zeman, 1963). These features support a monogenic disorder or an association of the holoprosencephaly sequence with other disorders.

Secondly, familial cases of holoprosencephaly (also known as non-syndromic holoprosencephaly cases) have X-linked, and autosomal dominant and recessive modes of inheritance (Cohen, 1989a,b). The autosomal dominant and recessive modes of inheritance are difficult to ascertain due to the incomplete penetrance of the holoprosencephaly phenotype. Several studies suggest a 70-82% penetrance for autosomal dominant inheritance of holoprosencephaly (Cohen, 1989a,b; Odent et al., 1998). The recurrence risk of autosomal dominant HPE is 29-35%, while in sporadic cases, 6% (Peebles, 1998). For recessive HPE cases, the recurrence risk is as high as 25% (Cohen and Sulik, 1992). There is no correlation observed between extracranial facial findings and race. However, males with autosomal dominant mutations are proposed to have a greater risk for major malformations outside the central nervous system than do females (Suthers et al., 1999). A possible reason for this is that the reproductive fitness of males is lower than females, thus allowing the trait greater penetrance in males (Suthers et al., 1999). Of noteworthy interest, subjects affected with autosomal recessive holoprosencephaly tend to have more severe features within the holoprosencephaly spectrum than autosomal dominant cases (Cohen, 1989a,b). A possible reason is the incomplete penetrance for autosomal dominant HPE.

Thirdly, monozygotic twins demonstrate greater concordance for holoprosencephaly features than do dizygotic twins (Cohen, 1989a,b), supporting a greater contribution of the genetic background to the pathogenesis of holoprosencephaly rather than environmental insults.

Fourthly, a collection of chromosomal holoprosencephaly cases, also known as syndromic holoprosencephaly cases, has been reported (Cohen, 1989a,b; Cohen and Sulik, 1992). Between 24 and 45% of HPE cases are chromosomal in origin, with trisomy 13 occurring about 70% of the time (Cohen and Sulik, 1992). Aneuploidy (such as rings, duplications, deletions and translocations) and triploidy is associated with holoprosencephaly. Syndromic HPE subjects also demonstrate more severe clinical findings, with extracranial anomalies present in conjunction with the typical holoprosencephaly craniofacial sequence. This is likely due to multiple genes affected in the chromosome aberrations. If a parent has a karyotypic anomaly, the recurrence risk for HPE is 1% (Peebles, 1998). It is higher if both parents have a karyotypic anomaly. No correlation between race or sex with chromosomal or non-chromosomal incidence of holoprosencephaly has yet been established (Whiteford and Tolmie, 1997; Olsen et al., 1997; Moog et al., 2001; Croen et al., 1996; Cohen, 1989a,b).

1.12. MOLECULAR PATHWAYS AND HOLOPROSENCEPHALY GENES

1.12.1. Definition of holoprosencephaly loci

Based on linkage analyses and on the alignment of chromosome aberrations, 12 holoprosencephaly loci are postulated on 11 chromosomes (Table 1-1) (Roessler and

Table 1-1: Proposed genome wide holoprosencephaly loci and genes

HPE locus	Map position	Gene studied	Putative function	Type of HPE	Mutations found in HPE patients
HPE 1	21q22.3	Lanosterol synthetase (*LS*)	Cholesterol synthesis	Recessive	Yes
HPE 2	2p21	Sine oculis 3 (*SIX3*)	Transcription factor	Dominant	Yes
HPE 3	7q36	Sonic hedgehog (*SHH*)	Developmental ligand	Dominant	Yes
HPE 4	18p11.3	TG interacting factor (*TGIF*)	Transcription factor	Dominant	Yes
HPE 5	13q32	Zinc finger protein in cerebellum 2 (*ZIC2*)	Transcription factor	Dominant	Yes
HPE 6	3p24-pter				
HPE 7	13q12-q13				
HPE 8	14q13	This thesis			
HPE 9	20p13				
HPE 10	1q42-qter				
HPE 11	5p				
HPE 12	6q26-qter				
Locus name not given	9q22.3	Patched 1 receptor (*PTCH*)	Receptor for sonic hedgehog ligand	Dominant	Yes
	10q11.2	Dickkopf 1 (*DKK1*)	Developmental ligand	Dominant	Yes
	3p21-p23	Tetracarcinoma derived growth factor 1 (*TDGF1*)	Developmental ligand and co-receptor	Dominant	Yes

See text for details and references.

Muenke, 1998). Chromosome 13 has two HPE loci located at the proximal and distal cytogenetic intervals. The minimal critical regions for most holoprosencephaly loci have been proposed mainly through the analyses of disease associated chromosome aberrations. However, non-syndromic holoprosencephaly cases were useful in linkage analyses to define the minimal critical region for HPE3, the first characterized locus, between the markers *D7S550* and *D7S22* at 7q36 (Muenke et al., 1994; Gurrieri et al., 1993). This region was confirmed with the mapping of a t(7;9)(q36;q34) HPE patient, who demonstrated position-effect on the *SHH* gene. Genotyping of 125 affected cases from eight of nine families with autosomal dominant HPE showed linkage to the marker *D7S22*, suggestive of evidence for genetic heterogeneity. Mutations were eventually found within the *SHH* gene at 7q36 (Roessler et al., 1996, 1997; Orioli et al., 2001, 2002; Heussler et al., 2002; Nanni et al., 1999, 2001; Kato et al., 2000). The next holoprosencephaly loci to be defined were, HPE1 at 21q22.3; HPE2 at 2p21, HPE4 at 18p11.3 and HPE5 at 13q32, by the mapping of interstitial deletion and translocation cases (Roessler and Muenke, 1998). The HPE loci 6 to 12 on chromosomes 13,3,14,20,1,5 and 6 were proposed by the alignment of disease associated chromosome aberrations using syndromic HPE cases.

It is noteworthy that disease associated deletion and translocation chromosomes may not always result in a phenotype based on disruption or deletion of the coding region of a gene. Deletion or translocation breakpoint intervals can mediate position-effect on neighboring gene(s) by altering correct gene transcriptional expression via changes in the promoter, enhancer/silencer or local chromatin structure (Kleinjan and van Heyningen, 1998). Position effect can be mediated from the 5' or 3' end of a gene and may act as far as 1 Mb. Examples of such phenomena have been observed with the *PAX6* gene in Aniridia, *SOX9* gene in Campomelic dysplasia, and *SHFM1* gene of Split hand/split foot malformation (reviewed by Kleinjan and van Heyningen, 1998).

Mutational analyses have been performed on three genes, namely *PTCH, DKK1* and *TDGF1*, which map to undefined holoprosencephaly loci. Currently, investigations are determining whether the map positions of the latter genes correlate with chromosomal aberrations in patients with the holoprosencephaly spectrum. For all genes with identified mutations among holoprosencephaly patients, no phenotype correlations can be discerned, since there is inter and intrafamilial variability in phenotype with the same mutation (Wallis and Muenke, 2000). A brief description of these genes and their involvement in developmental pathways, is outlined in the subsequent sections and Figure 1-2.

1.12.2. Sonic hedgehog signaling pathway

Members of the hedgehog signaling pathway were initially identified by a screen for mutations affecting *Drosophila* embryonic and larval development (Nusslein-Volhard and Wieschaus, 1980). Three vertebrate homologues of the *Drosophila* hedgehog gene, namely Indian, desert and sonic were later identified (Hammerschmidt et al., 1997). Several reviews have extensively described the details of this pathway (Villavicencio et al., 2000; Ingham and McMahon, 2001; Briscoe and Ericson, 1999; Hammerschmidt et al., 1997). The generic sonic hedgehog pathway essentially consists of a ligand, extracellular spatial localizers of the ligand, transmembrane receptors, co-ligand modulators, intracellular signal transducers and transcription factors. The pathway has been associated with many congenital disorders such as Gorlin syndrome, Pallister-Hall

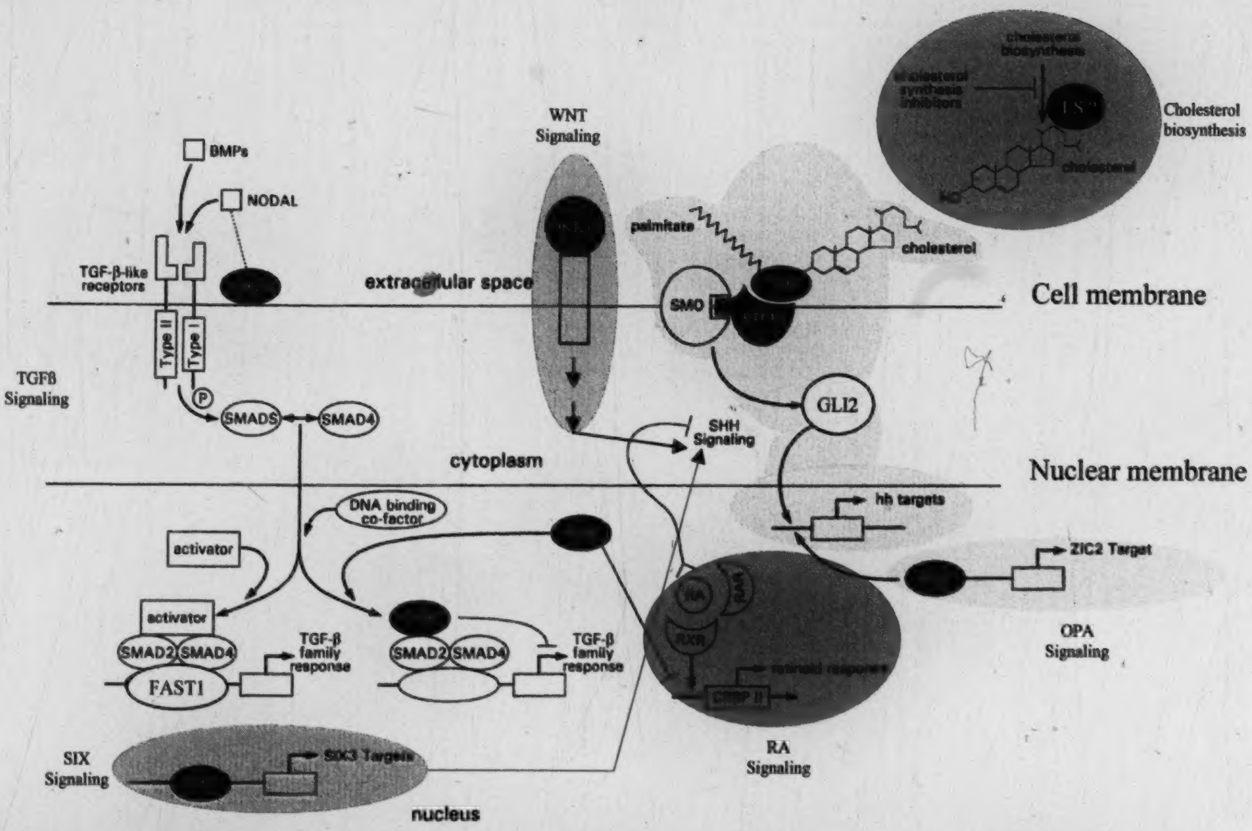

BMPs
NODAL
TGF-β-like receptors
Type II
Type I
P
TGFB Signaling
SMADS
SMAD4
extracellular space
WNT Signaling
palmitate
cholesterol
SMO
GLI2
SHH Signaling
cholesterol biosynthesis
cholesterol synthesis inhibitors
cholesterol
HO
Cholesterol biosynthesis
Cell membrane
cytoplasm
Nuclear membrane
DNA binding co-factor
activator
activator
SMAD2
SMAD4
FAST1
TGF-β family response
SMAD2
SMAD4
TGF-β family response
hh targets
ZIC2 Target
OPA Signaling
RA
RAR
RXR
retinoid response
CRBP II
RA Signaling
SIX3 Targets
SIX Signaling
nucleus

Figure 1-2: Possible interactions among genetic pathways associated with holoprosencephaly. Each genetic pathway is highlighted in a different color. These interactions are only suggestive and arise from experimental findings found among several organisms. Genes mutated among holoprosencephaly patients are shown in dark ovals. See text for details on each pathway. Modified from Ming and Muenke, 2002 with permission to use Figure 1 from The University of Chicago Press.

syndrome, postaxial polydactyly, Saethre-Chotzen syndrome, Rubinstein-Taybi syndrome, and several carcinomas (reviewed by Briscoe and Ericson, 1999).

1.12.2.1. Sonic Hedgehog gene (*SHH*)

Extensive studies in mice, zebrafish and *Xenopus* have shown *Shh* involvement in angiogenesis, haematopoiesis, and the formation of the cerebellum, eye, gut, hair, heart, limb, lung, muscle, neuron, oligodendrocyte, pancreas, prostate, tooth and tongue (Ingham and McMahon, 2001). *Shh* is expressed in a gradient in the notochord and ventral floor plate to pattern the developing neural tube and limbs (Ingham and McMahon, 2001). *Shh* is also involved in establishing the anterior-posterior and left-right body axes (Ingham and McMahon, 2001). Sonic hedgehog (*SHH*) was the first human HPE gene to be discovered (Belloni et al., 1996; Roessler et al., 1997). The gene contains three exons that encode a 462 amino acid protein with high conservation among vertebrates (Marigo et al., 1995). Expression studies in vertebrates have demonstrated *Shh* to be expressed in the Hensen's node, notochord, ventral midline of the prosencephalon and in the limb bud zone of polarizing activity (Roelink et al., 1995). Moreover, *Shh* triggers the differentiation of the floor plate and ventral neurons throughout the neural tube and alters the expression of numerous developmental genes (Chiang et al., 1996). Developmental expression studies on 26 to 37 day old human embryos demonstrated *SHH* expression at day 26 in the notochord, entire floorplate, spinal cord, hindbrain and forebrain (Odent et al., 1999). By 37 days of gestation, *SHH* is expressed in the entire central nervous system, notochord, floorplate, foregut, hindgut, lung buds and in the posterior mesenchyme of the fore and hind limb buds (Odent et al., 1999). The majority of these tissues are affected in patients with HPE at 7q36. The *Shh* gene has been disrupted in mice. Mice homozygous for the disrupted *Shh* gene are still born pups affected with cyclopia, CNS midline anomalies, lack of ventral forebrain development, and have defective limb, heart, lung, kidney and foregut (Chiang et al. 1996). An assortment of human mutations of many types have been found throughout the *SHH* gene in about 17% of familial HPE and 3.7% of sporadic HPE cases (Roessler et al., 1996; Kato et al., 2000; Roessler et al., 1997; Nanni et al., 1999). It is also interesting that for familial HPE cases, only 37% with structural defects have mutations in *SHH*. No mutations were found in a panel of patients with isolated midline, lateral or bilateral clefting (Orioli et al., 2002).

The sonic hedgehog protein is a secreted ligand that undergoes autoproteolytic cleavage into an amino terminus 19 kD domain and a carboxyl terminus 25 kD domain (Perler, 1998). The 25 kD domain has cholesterol transferase activity which adds a cholesterol moiety to the 19 kD domain (Perler, 1998). The 19 kD domain acts as the functional ligand and is believed to require a cholesterol moiety for its spatial restriction to the site of activity. The 19 kD domain has been found to also have a palmitate moiety (Pepinsky et al., 1998). The role of palmitate is unknown, however it has been demonstrated to affect the activity of the ligand. The fact that the sonic hedgehog protein requires cholesterol for its function explains why HPE features are observed in 2-4% of patients affected with Smith Lemli Optiz syndrome, a cholesterol biosynthesis disorder (Roessler and Muenke, 1998).

1.12.2.2. Patched-1 receptor gene (*PTCH*)

The *PTCH* gene encodes a receptor for the SHH ligand (Villavicencio et al., 2000). The binding of the SHH ligand to the PTCH receptor releases the PTCH induced inhibition of another receptor SMOOTHENED (Villavicencio et al., 2000). Expression studies in mice demonstrate the *Ptch* gene has overlapping spatial and temporal expression with the *Shh* gene (Platt et al., 1997). Loss of function mutations in the *PTCH* gene are implicated in basal cell nevus syndrome (Gorlin syndrome), where *PTCH* acts both as a developmental gene and as a tumor suppressor gene (Hahn et al., 1999). The same phenomenon is seen in heterozygous *Ptch* knockout mice that survive, but display skeletal abnormalities, neural tube defects, generalized overgrowth and basal cell nevus carcinoma (Goodrich et al., 1997; Hahn et al., 1998). Homozygous *Ptch* knockout mice die before E9.0 and show defective neural tube patterning (Goodrich et al., 1997). Interestingly, a conditional gain of function of *Ptch* expression in the developing mouse, results in fusion of the lateral ventricles as seen in holoprosencephaly (Goodrich et al., 1999). Several missense mutations within the intracellular and extracellular loops have been found in a panel of HPE patients (Ming et al., 2002). These mutations are proposed to exert gain of repressive functions in the PTCH receptor that abrogate the SHH mediated signaling (Ming et al., 2002).

1.12.3. Cholesterol biosynthesis

Cholesterol has many biological roles (reviewed by Salen et al., 1996). The link between cholesterol biosynthesis and holoprosencephaly stems from the fact that cholesterol limits the spatial regions of SHH signaling and is required to assist the transduction of the signal within target cells. Other evidence comes from the fact that pregnant rats treated with cholesterol biosynthesis inhibitors such as triparanol, AY9944 or BM 15.766 cause holoprosencephaly (Gofflot et al., 2001), by downregulating the SHH signaling pathway in the developing central nervous system during embryogenesis. A homozygous knock out of the murine *Megalin* gene results in holoprosencephaly with pups dying perinatally of respiratory distress (Willnow et al., 1996). *Megalin* is a member of the low density lipoprotein receptor gene family that transports cholesterol into the developing neuroepithelium (Willnow et al., 1996).

1.12.3.1. Lanosterol synthetase gene (*LS*)

Lanosterol synthetase (*LS*), a 22 exon containing gene, catalyzes the cyclization step in cholesterol biosynthesis (Roessler et al., 1999). Mutational analysis on three unrelated HPE patients with 21q22.3 hemizygous deletions, and 20 unrelated families with autosomal recessive mode of inheritance for HPE failed to show any mutations that have a phenotypic consequence. Since cholesterol also acts as a substrate for the synthesis of sterols and hormones and maintains the integrity of cell membranes (Roessler et al., 1999), homozygous or compound heterozygous mutations may be lethal.

1.12.4. Odd paired signaling pathway

The odd-paired (opa) family of transcription factors was initially found to serve in the parasegmental subdivision of the *Drosophila* embryo by controlling the expression of genes such as wingless (*Wg*) and engrailed (*En*) (Bhat, 1999; Benedyk et al., 1994). *Opa*, *Wg* and *En* genes all function in the hedgehog signaling pathway (Benedyk et al., 1994; Rohr et al., 1999). The odd-pair transcription factor family has a zinc finger domain and the vertebrate homologues thus far are ZIC1, 2, 3, 4 and 5 (Aruga et al., 1996; Nakata et al., 2000). The mouse knock-out of the *Zic1* gene results in skeletal, cerebellar and behavioral problems (Ogura et al., 1998), while the human *ZIC3* gene is mutated in patients with situs inversus and situs ambiguous (Gebbia et al., 1997).

1.12.4.1 Zinc finger protein in cerebellum 2 gene (*ZIC2*)

During mouse embryogenesis, *Zic2* is expressed as early as E7 in the gastrula, then in the dorsal neural tube, neural retina of the developing eye and distal limb mesoderm under the apical ectodermal ridge and precartilaginous condensations (Nagai et al., 1997). In older murine fetuses and adults, expression is restricted to the cerebellum (Nagai et al., 1997). *Zic2* acts as a pre-pattern gene to synergize with glioma 3 (*Gli3*) downstream of the Shh signaling pathway (Koyabu et al., 2001; Mizugishi et al., 2001). The Zic2 protein also binds to the dopamine 1a receptor (Yang et al., 2000) and apolipoprotein E (Salero et al., 2001) promoters, however with unknown biological consequences. A homozygous knock down of the murine *Zic2* gene results in neural tube defects and holoprosencephaly (Nagai et al., 2000). Heterozygous *Zic2* knock-down mice have no obvious central nervous system structural abnormalities except abnormal behaviors reminiscent of the schizophrenia spectrum (Ogura et al., 2001). Specifically, these mice demonstrated reduced acoustic startle response and prepulse inhibition of the acoustic startle response, that are typically observed in neuropsychiatric disorders like schizophrenia and schizotypal personality disorder (Ogura et al., 2001). In humans, the *ZIC2* gene has three exons that encode 3.2 and 3.5 kb alternative transcripts which are expressed in the fetal brain (Brown et al., 1998). Screening a panel of HPE patients demonstrated mutations in the *ZIC2* gene of several sporadic and familial HPE cases (Oriolli et al., 2001; Brown et al., 2001; Brown et al., 1998). Six cases had expansion of a polyalanine tract. All patients with mutations had normal or mild midface hypoplasia and none had single central maxillary incisor, consistent with the numerous cases of trisomy 13 that are diagnosed with holoprosencephaly and mild facial signs (Brown et al., 2001). One patient had a normal face and middle interhemispheric variant of holoprosencephaly (Brown et al., 2001). No mutations of *ZIC2* were found in a panel of patients with neural tube defects (Brown et al., 2002).

1.12.5. Sine oculis pathway

The sine oculis pathway is comprised of genes that are homologous to the *Drosophila* sine oculis (*So*) gene and are primarily involved in the patterning of the eye in conjunction with other genes including eyeless (*Ey*), dachshund (*Dach*) and those of the hedgehog signaling pathway (Kawakami et al., 2000). Members of this pathway have been associated with bilateral anopthalmia, pituitary abnormalities, and myotonic dystrophy (Kawakami et al., 2000).

1.12.5.1. Sine oculis 3 gene (*SIX3*)

The vertebrate *Six3* gene has been demonstrated to be involved in the patterning of the eye and midline of the forebrain in mouse, fish and zebrafish (Oliver et al., 1995, 1996; Kobayashi et al., 1998). *Six3* is expressed as early as E6.5 in the rostral anterior region of the neural plate, midline of ventral forebrain, optic vesicles, optic chiasma, nasal placodes and Rathke's pouch, up to E11.5 (Oliver et al., 1995; Kawakami et al., 1996). Expression becomes restricted to the eyes and parts of the nasal cavities by E18 (Oliver et al., 1995). Over expression of *Six3* in zebrafish results in enlargement of the rostral forebrain and optic stalk (Kobayashi et al., 1998), while in medka fish, it results in ectopic retina formation (Loosli et al., 1999). In both cases, empty spiracles homolog 2 (*Emx2*) and paired box 2 (*Pax2*) genes are induced, while the orthodenticle homolog 2 (*Otx2*) and paired box 6 (*Pax6*) genes have been proposed to control *Six3* expression (Rhinn et al., 1998; Chow et al., 1999). The *SIX3* gene, has three exons, with the second exon encoding the SIX domain (115 amino acids) and homeodomain (60 amino acids) (Leppert et al., 1999). No mutations were found in patients with septo-optic dysplasia (Leppert et al., 1999), however by screening a panel of HPE patients, several mutations were found in sporadic and familial HPE cases (Wallis et al., 1999; Pasquier et al., 2000).

1.12.6. Transforming growth factor beta signaling pathway

The TGFß pathway has been extensively studied for its role in establishing the anterior-posterior and left-right axes in the developing vertebrate embryo (reviewed by Schier and Shen, 1999; Goumans and Mummery, 2000; Hill, 2001). Signaling is mediated by activin-like receptors, ligands with an EGF-CFC (epidermal growth factor-like; Cripto, Frl-1, Cryptic) domain as well as those of the lefty and cerebrus families, and transcription factors that modulate the SMAD (mothers against decapentaplegic) family of transcription factors (Schier and Shen, 1999; Goumans and Mummery, 2000). The type II receptor is a serine/threonine kinase that phosphorylates the type I receptor which transduces the signal (Weinstein et al., 2000). This pathway acts before gastrulation commences and is responsible for angiogenesis, organogenesis, immune system function and wound healing (Weinstein et al., 2000). A variety of embryopathies, carcinomas and immune dysfunctions are linked to defects of the TGFß signaling pathway (Weinstein et al., 2000; Marek et al., 2002; Rooke and Crosier, 2001).

1.12.6.1. TG interacting factor gene (*TGIF*)

The TG-interacting factor (*TGIF*) encodes a homeodomain protein that was initially isolated by its ability to compete with retinoid X receptor in the binding to the retinoid receptor response elements (RXR), thereby causing repression (Bertilino et al., 1995). The TGIF protein also competes with p300 histone acetylases in the competitive binding to the SMAD2 complex (Luo et al., 1999). TGIF can also interact with SIN3A or c-Jun corepressors and recruit histone deacetylases via interaction with its carboxyl terminal domain to function in the c-Jun N terminal kinase (JNK) or steroid hormone response pathways (Melhuish et al., 2001; Wotton et al., 2001; Pessah et al., 2001). The interaction of TGIF with EGF (epidermal growth factor) is thought to stabilize the TGIF protein via phosphorylation to function in the Ras-Mek pathway (Lo et al., 2001; Wotton and Massague, 2001). Expression studies demonstrate that *Tgif* is expressed in the forebrain and craniofacial structures of E15 mouse embryos (Gripp et al. 2000) but later is expressed predominantly in the central nervous system (Bertolino et al., 1995). Heterozygous mutations have been found in the coding regions of *TGIF* from mutational analysis of HPE patients (Gripp et al. 2000; Chen et al., 2002).

1.12.6.2. Teratocarcinoma derived growth factor 1 gene (*TDGF1*)

The TDGF1 gene, also known as CRIPTO, is a member of the EGF-CFC family of ligands, including one-eyed pinhead (Oep). The TDGF1 protein has an epidermal growth factor (EGF)-like motif (which is O-linked fucose monosaccharide modified) and a CFC domain which acts as a co-receptor and co-ligand for the binding of nodal (Yan et al., 2002; Yeo and Whitman, 2001). Specifically, the CFC domain permits Tdgfl to bind to the activin-like type I serine/threonine kinase receptor and the EGF-like motif facilitates Nodal to bind to the Activin-like type II receptor (Yan et al., 2002; Yeo and Whitman, 2001). The Tdgfl ligand can rescue the Oep zebrafish phenotype (Gritsman et al., 1999) which has the holoprosencephaly spectrum. Expression of the murine *Tdgfl* gene is seen in the primitive streak and head process (Ding et al., 1998), but later in the heart (Dono et al., 1993). Homozygous knock out *Tdgfl* mice die before 10.5 dpc, with absence of the primitive streak, abnormal anterior-posterior axis formation and absence of embryonic mesoderm and endoderm (Ding et al., 1998). Tdgf1 is also thought to orient the anterior-posterior axis, can activate the Ras/Raf/MAPK pathway in cultured mammary epithelial cells (Kannan et al., 1997), and is possibly involved in the formation of many types of tumors (Admamson et al., 2002). The human *TDGF1* gene, which has six exons, is expressed in the adult colon, and various carcinomas (gastric, pancreatic, lung and breast) (Minchiotti et al., 2002). A missense mutation was found in the CFC domain of a patient with lobar holoprosencephaly or possibly middle interhemispheric variant of HPE (de la Cruz et al., 2002). This mutation resulted in a non-functional protein (de la Cruz et al., 2002). Another uncommon conservative missense mutation was found in the carboxyl terminus in a patient with semilobar holoprosencephaly (de la Cruz et al., 2002).

1.12.7. WNT signaling pathway

The Wnt signaling pathway is involved in the development of the limbs, brain, kidney and reproductive tract, and also in the genesis of many types of cancers. Details of this pathway are documented by several groups (Talpale and Beachy, 2001; Smalley and Dale, 1999).

1.12.7.1. Dickkopf 1 gene (*DKK1*)

The Dickkopf 1 (*DKK1*) gene encodes a secreted ligand with two cysteine rich domains. The DKK1 protein binds to the lipoprotein receptors 5 and 6, and Kremen receptors 1 and 2 to inhibit Wnt signaling (Mao et al., 2001,2002; Brott and Sokol, 2002). A homozygous knock out of the murine *Dkk1* gene results in the loss of forebrain and in limb digit malformation (Mukhopadhyay et al., 2001), while in *Xenopus*, loss of protein activity results in microcephaly and cyclopia (Niehrs, et al., 2001), thus showing its role in forebrain specification and axial mesendoderm patterning. Screening of the coding regions of the human *DKK1* gene, that has four exons, led to identifying missense mutations in HPE families (Roessler et al., 2000). However, functional assays of these mutations did not demonstrate an affect on protein function (Roessler et al., 2000).

1.12.8. Retinoic acid signaling pathway

Retinoic acid is a tetratogen known to produce holoprosencephaly in addition to many other embryopathies, hematological malignancies and many types of carcinomas. Details of this pathway are documented by several groups (Chambon, 1993; Mandelli et al., 2002; Sun and Lotan, 2002). Knock out studies of the retinoic acid receptor and retinoid X receptors of this signaling pathway demonstrate the role of retinoic acid in patterning the mesectoderm (Pauken et al., 1999; Mark et al., 1998). Furthermore, the expression of sonic hedgehog was absent in craniofacial primordia of chick when exposed to high doses of retinoic acid (Helms et al., 1997). Given this, retinoic acid is proposed to obscure the SHH signaling pathway in the development of the frontonasal and maxillary processes thus resulting sometimes in the holoprosencephaly phenotype (Helms et al., 1997). Mutational analyses have not yet been reported for any key members of the retinoic acid signaling pathway.

1.12.9. Mutations in HPE patients are sometimes found in more than one gene

Mutation studies have exhibited heterozygous mutations in the *SHH* gene in conjunction with other HPE genes such as *ZIC2* and *TGIF*, in HPE affected subjects (Ming and Muenke, 2002), suggesting that the penetrance and expression of the HPE phenotype is dependent on the alleles of genes involved in forebrain and craniofacial formation. No correlations can thus far be found between the HPE phenotype severity and having mutations in more than one of the HPE genes. It is likely that gene-environment interactions also contribute to the penetrance and expression of the HPE phenotype. Furthermore, mutation analyses on 200 to 500 affected HPE subjects have demonstrated only 15 to 20% have heterozygous mutations in the HPE genes listed in Table 1-1 (Nanni et al., 2000, Ming and Muenke, 2002). This supports the notion that many HPE genes are yet unidentified.

1.12.10. Proposed candidate genes for holoprosencephaly

Candidate genes should include those that have roles in forebrain patterning; cholesterol biosynthesis; formation of the prechordal plate, anterior notochord or anterior neural tube; and are involved in the sonic hedgehog, retinoic acid, odd-paired, transforming growth factor beta and/or *sine oculis* pathways. Moreover, animal models such as zebrafish, with genes like cyclops and squint (Rebaglianti et al., 1998; Feldman and Stemple, 2001; Gritsman et al., 1999) have yet to be characterized human orthologs that could potentially cause the HPE phenotype. Genes that cause HPE neomorphic phenotype, such as *Bmp4* (bone morphogenetic protein 4) are also candidates for forebrain and craniofacial morphogenesis (Golden et al., 1999). A list of potential HPE candidate genes is reported in several reviews (Roessler and Muenke, 1999; Wallis and Muenke, 1999; Muenke and Beachy, 2000).

1.13. MOUSE AS A MODEL ORGANISM TO STUDY HOLOPROSENCEPHALY

The mouse has served as an excellent model organism to study mammalian developmental neurogenetics at low cost, due to its small size, resistance to infection, large litter size and short onset generation period (Hogan et al., 1994). In this dissertation, the C56BL/6J mouse strain was used for embryonic expression studies. This strain of mice was one of the first few that initially had the neuroanatomy and embryonic theiler staging deciphered (Miyake et al., 1996). It has been used in extensive studies by many researchers world wide to study the embryonic expression pattern of genes during embryonic, neonatal and adult development. Furthermore, the neuroembryology and neuroanatomy are very similar to human in that it contains a corpus callosum which certain mice strains like Balb/c and 129 strains lack 20% of the time (Livy et al., 1997). The comparative mouse-human embryonic period that is required for craniofacial development is as follows:

Structure	C57BL/6J Mouse		Human	
Primitive streak	6.5 to 7.5	dpc	13 to 15	dpc
Central nervous system architecture	7.5 to 11.5	dpc	18 to 43	dpc
Face	> 8.5 to <14	dpc	22 to 70	dpc
Corpus callosum	9 to 18	dpc	18 to 140	dpc

1.14. PREAMBLE TO THE STUDIES OF THIS THESIS AND PRESENTATION OF THE GENERAL HYPOTHESIS

Both linkage analyses and the study of chromosome rearrangements segregating in subjects with holoprosencephaly have suggested that 1) there is a genetic basis for holoprosencephaly, and 2) there is tremendous genetic heterogeneity of the holoprosencephaly trait. When this thesis project started in the winter of 1998, there was evidence for a human chromosome 14 association with holoprosencephaly, based on reported non-random cytogenetic findings of patients with a chromosome 14 aberration and a phenotype within the holoprosencephaly spectrum. Four such cases involved deletions of the proximal region of human chromosome 14 in sporadic cases of alobar, semilobar and lobar holoprosencephaly (Levin and Surana, 1991; Bruyere et al., 1996; Chen et al., 1997; Jules LeRoy (personal communication)). One Robertsonian translocation (13q14q) case was reported (Chow et al., 1996), however with uncertainty about whether the suspected additional unbalanced rearrangement involved chromosome 13 or 14 as the cause of holoprosencephaly. The locus on chromosome 14q13, was designated as HPE8 by Roessler and Muenke (1999) (Figure 1-3). Before this research had started, there were also two reported cases of patients with large proximal chromosome 14q deletions (Shapira et al., 1994) showing features within the mild holoprosencephaly spectrum involving hypoplasia of the corpus callosum, and one translocation (4;14) patient with agenesis of the corpus callosum (Maximillian Muenke, personal communication). Five other cases with features within the mild holoprosencephaly spectrum were later reported during the course of this research (Ramelli et al., 2000; also reported herein). This suggested two phenomena. Firstly, the proximal region of chromosome 14q contained a locus for holoprosencephaly and another locus for acallosal defects. Secondly, patients with classic holoprosencephaly or features within the mild holoprosencephaly spectrum represented variable expression of the holoprosencephaly phenotype, suggesting that a single locus caused both spectra of phenotypes.

The studies described in this dissertation were formulated to test the following general hypothesis:

The proximal region of human chromosome 14, namely, 14q13, contains one or more genes that when mutated cause the holoprosencephaly spectrum of phenotypes, including severe alobar holoprosencephaly to mild holoprosencephaly features.

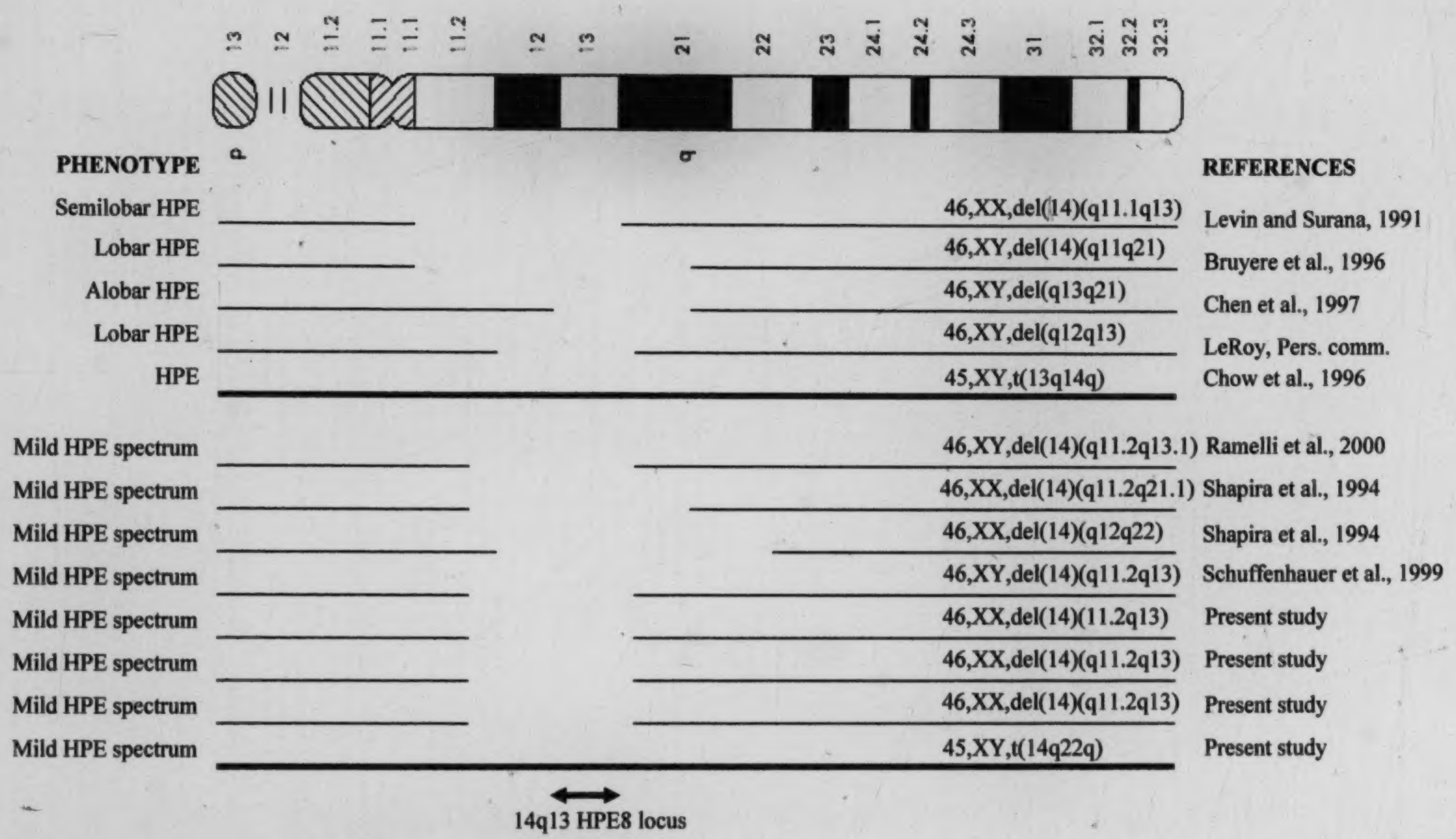

13
12
11.2
11.1
11.1
11.2
12
13
21
22
23
24.1
24.2
24.3
31
32.1
32.2
32.3
p
q
PHENOTYPE
REFERENCES
Semilobar HPE
46,XX,del(14)(q11.1q13)
Levin and Surana, 1991
Lobar HPE
46,XY,del(14)(q11q21)
Bruyere et al., 1996
Alobar HPE
46,XY,del(q13q21)
Chen et al., 1997
Lobar HPE
46,XY,del(q12q13)
LeRoy, Pers. comm.
HPE
45,XY,t(13q14q)
Chow et al., 1996
Mild HPE spectrum
46,XY,del(14)(q11.2q13.1)
Ramelli et al., 2000
Mild HPE spectrum
46,XX,del(14)(q11.2q21.1)
Shapira et al., 1994
Mild HPE spectrum
46,XX,del(14)(q12q22)
Shapira et al., 1994
Mild HPE spectrum
46,XY,del(14)(q11.2q13)
Schuffenhauer et al., 1999
Mild HPE spectrum
46,XX,del(14)(11.2q13)
Present study
Mild HPE spectrum
46,XX,del(14)(q11.2q13)
Present study
Mild HPE spectrum
46,XX,del(14)(q11.2q13)
Present study
Mild HPE spectrum
45,XY,t(14q22q)
Present study
14q13 HPE8 locus

Figure 1-3: Alignment of the cytogenetic intervals of chromosome 14 aberration cases associated with the holoprosencephaly spectrum. Broad bars are Robertsonian translocation cases.

The final goal of this thesis was to suggest one or more candidate genes for holoprosencephaly at the characterized locus at chromosome 14q13. To achieve this, the specific aims were to:

1. To collect DNA or cell line specimens from a panel of patients with proximal chromosome 14q rearrangements and use these patients to molecularly define a holoprosencephaly minimal critical region at 14q13;
2. To construct physical contigs of the defined holoprosencephaly critical region;
3. To construct a transcript map of the minimal critical region and to identify the cDNA sequences and genomic structure of potential candidate genes for holoprosencephaly;
4. To ascertain possible functions of selected candidate genes for holoprosencephaly by means of bioinformatic analyses, human expression pattern studies, human cellular localization studies and developmental expression pattern studies using mouse as a model organism;
5. To determine whether a selected panel of translocation patients characterized in this research had disrupted or rearranged genes mapping to the proximal region of chromosome 14, in particular candidate genes for holoprosencephaly, or genes on the other derivative chromosome.

The studies described in this dissertation first outline the collection of DNA or cell lines of patients with translocations or deletions, of the proximal region of human chromosome 14q (Chapter Two). These patients either had mild to severe expression of the holoprosencephaly phenotype or completely lacked the holoprosencephaly spectrum of phenotypes. The subsequent mapping using microsatellite typing and analysis of flow sorted chromosomes led to the definition of a holoprosencephaly minimal critical region at chromosome 14q13 between *D14S49* and *D14S1014*. In the second series of studies (Chapter Three), YAC and BAC physical contigs were constructed in the minimal critical region, with an estimated size of two Mb. A transcript map of 34 potential genes was defined, five of which were proposed to be candidates for holoprosencephaly. The third series of studies (Chapter Four) reports further selection of potential candidate genes for holoprosencephaly and characterization of genes with respect to the cDNA sequence, genomic structure, and possible functions ascertained by bioinformatic analysis, human expression pattern profiles, cellular localization studies, and embryonic expression studies using the mouse as a model organism. With these findings, three excellent candidates for holoprosencephaly are proposed. The final studies (Chapter Five) determined whether three translocation patients had rearrangements or disruptions within genes mapping at 14q13, particularly the holoprosencephaly candidates, and those mapping to the other derivative chromosome. Such analyses were used to explain the cause of the phenotype observed in the selection of translocation patients. Two of the three patients were found to have a disruption in one of the HPE8 candidate genes.

CHAPTER TWO

CLINICAL DESCRIPTION AND MAPPING OF PATIENTS WITH PROXIMAL CHROMOSOME 14q REARRANGEMENTS

The following people and cell bank provided clinical information, cytogenetic information, and blood/cell line specimens from patients and/or parents for mapping:- Drs. Maximillan Muenke (National Institute of Health, USA), Elaine Zackai (Pennsylvania School of Medicine,USA),Walter Muir (University of Edinburgh, UK), Chih-Ping Chen (Mackay Memorial Hospital, China), James Lupski (Baylor College of Medicine, USA), Paola Grammatico (University of Rome "La Sapienza", Italy), Oliver Quarrell (Sheffield Children's Hospital, UK), Simone Schuffenhauer (Ludwig-Maximilians University Munich, Germany), Jennifer Gerritsen (Alberta Children Hospital, Canada), Ross McLeod (Alberta Children Hospital, Canada), John Tolmie (University of Glasgow, UK), Koenraad Devriendt (University of Leuven, Belgium), Laskhmi Mehta (North Shore Hospital, USA), and the European Collection of Cell Cultures (UK).

Claire Chapman from the Oxford Chromosome Abnormality database (UK) provided a list of proximal chromosome 14q constitutional and acquired cases.

Nancy Cracknell (Hospital for Sick Children, Toronto) established lymphoblastoid cell lines for selected patients and parents.

Patricia O'Brien with Dr Malcolm Ferguson-Smith (Cambridge University, UK) isolated normal and aberrant chromosomes of selected patients by flow cytometry.

The Cytogenetics facility (University of Alberta Hospital, Edmonton) performed karyotyping with G banding on selected family members of case 16.

Parts of this chapter have been published in:

1. **Kamnasaran D**, Gerritsen JA, McLeod DR and Cox DW. Features of the holoprosencephaly spectrum associated in sibs with a Robertsonian (14q;22q) translocation chromosome. Clinical Genetics 2001: 60(3):237-239.
2. **Kamnasaran D**, O'Brien PCM, Schuffenhauer S, Quarrell O, Lupski JR, Grammatico P, Ferguson-Smith MA and Cox DW. Defining the breakpoints of proximal chromosome 14q rearrangements in nine patients using flow sorted chromosomes. American Journal of Medical Genetics 2001: 201(2):173-182.
3. **Kamnasaran D** and Cox DW. Current Status of Human Chromosome 14. Journal of Medical Genetics 2002. 39:81-90.

2.1. INTRODUCTION

Isolated interstitial deletion and translocation cases involving the human chromosome 14q11.2-q21 cytogenetic interval are relatively rare, but at least 28 cases involving this region have been reported thus far (Das et al., 2002; Wang et al., 2001a; Kamnasaran et al., 2001; Buchanan et al., 1978; Ramelli et al., 2000; Breedveld et al., 2002; Krude et al., 2002; Bruyere et al., 1996; Chen et al., 1997, 1998; Cooke et al., 1989; Devriendt et al., 1998a; Govaerts et al., 1996; Kim et al., 1997; Iwantani et al., 2000; Levin and Surana 1991; Mehta et al., 1999; Grammatico et al., 1994; Schuffenhauer et al., 1999; Shapira et al., 1994). The molecular definition of chromosome rearrangements within this interval will contribute to identifying disease genes in addition to determining the mechanisms by which the rearrangements arise.

In this chapter, the definition of the deletion and translocation breakpoints of aberrant chromosomes from patients ascertained with and without the holoprosencephaly spectrum is described. The aberrant chromosomes from twelve patients were analyzed by flow sorted chromosome analysis, while an additional three patients with proximal chromosome 14q deletions were mapped by microsatellite typing. By using the mapping information obtained from the panel of patients, the ultimate goal of this chapter was to test the sub-hypothesis that "*a locus for holoprosencephaly exists at 14q13 (HPE8)*". Parental origin, sex, ethnicity and deletion sizes were also compared with the phenotype to ascertain any possible correlations. Furthermore, a family with two sibs affected with microforms of the holoprosencephaly spectrum was analyzed to determine whether the phenotype was associated with a 14q13 locus. Taken together, the findings of this chapter have excluded parts of chromosome 14q, and defined a locus for holoprosencephaly on human chromosome 14q13.

2.1.1. The association of genetic syndromes with proximal chromosome 14q

The proximal region of chromosome 14q contains loci for at least 15 disorders with nine identified causative genes (Kamnasaran and Cox, 2002; Wang et al., 2001; Kondoh et al., 2002). The disease genes at chromosome 14q13 include, *PAX9* (paired box 9) for oligodontia, *NKX2.1* (thyroid transcription factor 1) for benign hereditary chorea, *COCH* (cochlin) for dominant deafness 9, and *MIPOL1* (mirror image polydactyly 1) for tetramelic mirror image polydactyly (Kamnasaran and Cox, 2002; Kondoh et al., 2002; Breedveld et al., 2002). The use of positional cloning, including linkage and mutation studies, has primarily defined and identified these disease loci and genes. However, some loci and genes for diseases such as tetramelic mirror image polydactyly at 14q13 were identified by use of chromosome aberrations (Kim et al., 1997; Kondoh et al., 2002).

2.1.2. The association of psychiatric illnesses with proximal chromosome 14q

The quest to find genes responsible for psychiatric illnesses is a major challenge due to the complexity of the trait, improper clinical ascertainment and inadequate statistical methods. Regardless of these disadvantages, modest evidence has been suggested for bipolar disorder and schizophrenia at the proximal region of chromosome 14. Specifically, two bipolar loci at *D14S1280* and *D14S297*; and two schizophrenia loci at *D14S79* and *D14S306* have been proposed (Craddock and Lendon, 1999).

2.2. MATERIALS AND METHODS

2.2.1. Search for proximal 14q chromosome aberrations

Extensive searches on the databases of cytogenetic laboratories from North America and Europe were done to find unpublished human chromosome 14q proximal deletion and translocation cases. These databases included American Tissue Culture Collection (www.atcc.org), European Collection of Cell Cultures (www.camr.org.uk), Helix (www.helix.com), Oxford Chromosome Abnormality Database (www.hgmp.mrc.ac.uk/local-data/Cad_Start.html), Canadian cytogenetic laboratory databases (www.hrsrh.on.ca/genetics/canlabs.htm), Cancer Genome Anatomy Project (cgap.nci.nih.gov/ Chromosomes/Mitelman), and the Italian and Russian Cell line Collections (www.biotech.ist.unige.it/cldb/mcells.html). Previously published cases were also sought with Medline (www.ncbi.nlm.nih.gov) searches using the keywords – chromosome 14, corpus callosum, holoprosencephaly, deletion, translocation.

2.2.2. Flow sorting aberrant chromosomes from lymphoblast cell lines

Transformed lymphoblast cell lines were established at The Hospital for Sick Children (Toronto) by standard methods from 10 ml of blood from the patients (cases 7,11,14,16) and some of the parents (parents of case 14 and 16). Cell lines for cases 6,8,9 were obtained from the European Collection of Cell Cultures. Cell lines from cases 1,2,3,4,5,10,12 were obtained from collaborators. See Appendix 1 for details. Only DNA was available from cases 13 and 15. No specimen was available from case 17. Cell lines were cultured at 37°C in Roswell Park Memorial Institute medium (Gibco) plus 10% fetal bovine serum (Gibco), and were flow sorted as previously described (Ferguson-Smith, 1997). Three to five hundred copies of the normal or aberrant sorted chromosome 14 were subjected to DOP-PCR amplification before being used as a template for mapping (Telenius et al., 1992). Specifically, each 100 µl reaction contained: 0.2 mM dNTPs, 2 mM $MgCl_2$, 1 X PCR buffer (Perkin Elmer), 10 µg gelatin, 0.8 µM 6-MW primer (ccgactcgagnnnnnnatgtgg), 0.5% Triton-X 100, and 20 units of AmpliTaq (Perkin Elmer). The DOP-PCR cycle was: denaturation for 10 minutes, 15 cycles of 94°C-1 minute, 30°C - 1.5 minutes, 30 to 72°C at 0.2°/sec for 3 minutes, then briefly 72°C- 3 minutes extension, followed by 35 cycles of 94°C-1 minute, 54°C-1 minute and 72°C- 3 minutes. A final extension for 10 minutes, 72°C was done. 2 µl of DOP-PCR amplified flow sorted chromosomes were used as template.

2.2.3. Mapping aberrant breakpoints

Forty seven chromosome 14 specific polymorphic and non-polymorphic markers (Research Genetics, Huntsville, Alabama) were selected from the Whitehead Institute for Genome Research (www-genome.wi.mit.edu), CÉPH (www.cephb.fr), Généthon (www.genethon.fr) and Marshfield (www.marshmed.org/genetics) linkage maps in addition to the Whitehead Institute for Genome Research STS content YAC maps (www-genome.wi.mit.edu). The order of the markers listed in Table 2-1 from centromere to telomere, was ascertained from the genetic map positions, on the Whitehead Institute CÉPH YAC physical maps and any available human chromosome 14 genomic sequences (GenBank, www.ncbi.nlm.nih.gov). The DNA templates for each 20 µl PCR reaction included human genomic DNA, or flow sorted aberrant chromosomes. The PCR conditions and cycles used for each marker are reported in Appendix 2. Mapping of the aberrant chromosome breakpoints was done three to nine times for each marker.

2.2.4. Determining the parent of origin for deletions

Genomic DNA was isolated by standard methods from the lymphoblast cell lines of patients and parents or from blood of selected parents (Moore, 1999; Signer et al., 1988). Microsatellite markers with heterozygosity values of ≥ 0.3 were selected from the Whitehead Institute for Genome Research, CÉPH, Généthon and Marshfield linkage maps. Only those markers with orders ascertained the same on these maps and localized to the human chromosome 14q proximal region were used. Each PCR reaction included 40 ng genomic DNA from the mother, proband or father, 1, 1.5 or 2 mM $MgCl_2$, 1X PCR buffer (Perkin Elmer), 1 μM of each primer, 0.005 mM dATP, 0.05 mM dTTP, 0.05 mM dCTP, 0.05 mM dGTP, 0.2 μl ^{35}S α- dATP (37 Mbq, 1mCi), and 0.2 units of AmpliTaq (Perkin Elmer). The PCR conditions and cycles are stated in Appendix 2. The father's DNA specimen was unavailable for case six. The parental origin of the deleted chromosome for this case was determined by genotyping 40 ng of genomic DNA from the mother and proband, and 2 μl of DOP-PCR amplified flow sorted aberrant chromosomes of the proband, using conditions mentioned above and in Appendix 2.

2.2.5. Uniparental disomy analysis

Genomic DNA was isolated by standard methods from the lymphoblast cell lines established from case 16 and the parents of case 16. A panel of 17 chromosome 14q-qter and five chromosome 22q-qter microsatellite markers with heterozygosity values of > 0.3 were selected from the Whitehead Institute for Genome Research, CÉPH, Généthon and Marshfield linkage maps. Each PCR reaction included 40 ng genomic DNA from the mother, child or father, 1, 1.5 or 2 mM $MgCl_2$, 1x PCR buffer (Perkin Elmer), 1 μM of each primer, 0.005 mM dATP, 0.05 mM dTTP, 0.05 mM dCTP, 0.05 mM dGTP, 0.2 μl

^{35}S -α- dATP (37 Mbq, 1mCi), and 0.2 units of AmpliTaq (Perkin Elmer). The PCR conditions and cycles used for each marker are reported in Appendix 3.

2.2.6. Microdeletion analysis

A panel of 18 polymorphic and non-polymorphic markers mapping at 14q13 were selected from the Whitehead CÉPH YAC STS content map to search for microdeletions in case 16. The DNA template for each 20 µl PCR reaction included 40 ng genomic DNA, or 2 µl of DOP-PCR amplified flow sorted aberrant chromosomes. The PCR conditions and cycles used for each marker are reported in Appendix 2. Mapping of the aberrant chromosome breakpoints was done in triplicate for each marker.

2.3. CLINICAL DESCRIPTION OF PATIENTS USED IN MAPPING

Case 1 (Cell line L251DC)

This case, previously reported as patient 737 of Hou 251 (Shapira et al., 1994), is a Hispanic girl born at 39 weeks gestational age to her 37 year old gravida 2 para 2 mother, by cesarean section. Her birth weight was 2,455 g (10th centile), length was 48 cm (50th centile) and OFC measurement was 31 cm (-2 standard deviations). In her first two months of life she was treated for gastroesophageal reflux and renal tubular acidosis. She had poor weight gain, developmental delay and mildly dysmorphic face (depressed nasal bridge). An MRI examination showed hypoplasia of the corpus callosum. The family history is unremarkable. She has a 46,XX,del(14)(q11.2q21.1) karyotype.

Case 2 (Cell line L777DC)

This case, previously described as patient 777 of Hou 266 (Shapira, et al., 1994), is a Vietnamese girl born at 36 weeks gestational age to her 33 year old gravida 3 para 3 mother, by cesarean section. Her birth weight was 2,345 g (25th centile), length was 45.5

cm (25th centile) and OFC measurement was 32.5 cm (mean). She showed poor weight gain, developmental delay, hypotonia, delayed bone ossification, bilateral middle ear dysfunction, patent foramen ovale, patent ductus arteriosus and a mildly dysmorphic face. Examinations by MRI indicated asymmetry of the brain and calvaria with flattening of the left parietal region and prominence of the right, and thinning of the corpus callosum. She died at the age of 14 months from cardio-respiratory arrest. The family history is unremarkable. She had a 46,XX,del(14)(q12q22) karyotype.

Case 3 (Cell line L640DC)

This case, previously reported (Grammatico et al., 1994), is a Caucasian male born at 38 weeks gestational age. His birth weight was 3,270 g (75th centile), length was 52 cm (90 centile) and OFC measurement was 33 cm (mean). He had generalized hypotonia, bilateral cryptorchidism, left hip subluxation, dysmorphic thorax, a mildly dysmorphic face and autistic-like behavior. He has a 46,XY,del(14)(q11.2q13) karyotype.

Case 4 (Cell line HQ0004)

The proband was born at 36 weeks gestation to normal healthy Caucasian parents with a weight of 2, 962g (75th centile) and OFC measurement of 30.4cm (10th centile). During the neonatal period, she had some breathing and feeding difficulties. By the age of 3 months her growth parameters fell below the 3rd centile and she was hypotonic with some hyper-reflexia in the limbs. Cranial CT and MRI scans showed agenesis of the corpus callosum with delayed myelination in the cerebral cortex. At the age of 6 years she had severe psychomotor retardation, no eye contact, no speech, profound microcephaly with an OFC measurement of 43 cm, proportionate short stature, fluctuating tone, gastroesophageal reflux and sleep disturbance. She died during early childhood. She had a 46,XX,del(14)(q11.2q13) karyotype.

Case 5 (Cell line SAN)

This case, previously described (Schuffenhauer et al., 1999) was born to a healthy 33 year old Caucasian parent after an unremarkable pregnancy, with a weight of 2,500 g (10th centile), length of 47 cm (25th centile) and head circumference of 32 cm (10th centile). He had bilateral cleft lip, hypertelorism, iris coloboma, low set ears, feeding difficulties, severe psychomotor retardation, muscular hypotonia and persistent foramen ovale. Cranial MRI and CT scans demonstrated agenesis of the corpus callosum, asymmetric ventricles, and a reduced brain volume. He has a 46,XY,del(14)(11.2q13) karyotype.

Case 6 (Cell line MAA)

The proband was the first child of normal, healthy Caucasian parent with a weight of 2,080 g (-3 standard deviation), length of 46 cm (10th centile) and head circumference of 30 cm (-3 standard deviation). MRI examinations demonstrated agenesis of the corpus callosum with pachygyria and an underdeveloped pituitary gland. At age three months, she showed signs of microcephaly, prominent metopic suture, premature synostosis of the corona suture, and a mildly dysmorphic face with a short nose, long philtrum, flat broad nasal root and macroglossia. She had feeding difficulties, frequent hyperthermia, generalized muscular hypotonia, seizures, elevated thyroid stimulating hormone levels, sleep disturbance, hyperopia, nystagmus, psychomotor retardation with no eye contact and frequent pulmonary, gastrointestinal and genitourinary tract infections. Cytogenetic studies showed a 46,XX,del(14)(q11.2q13) karyotype.

Case 7 (Cell line L6873)

This case, previously reported (Cooke et al., 1989), was an adult Caucasian male with a height within the 10th centile, weight over the 97th centile and head circumference over the 95th centile. He had an underdeveloped phallus, undescended right testis and inguinal hernia, and demonstrated mild mental impairment. Cytogenetic studies showed a 46,XY,t(4;14)(q12;q13) karyotype. He had one son who had similar clinical findings and was a carrier of the translocation chromosome. The spouse was clinically normal and had a normal karyotype. No other family members were available for clinical and cytogenetic analyses.

Case 8 (Cell line B01060)

The proband had short stature only. No other structural or neurological defects were evident. Cytogenetic studies showed a 46,XX,t(4;14)(q21;q13) karyotype. No other family members were available for clinical and cytogenetic analyses.

Case 9 (Cell line BV0416)

The proband had mesomelic bone dysplasia and no other structural or neurological abnormalities. Both parents are clinically normal, and there is no family history of mesomelic bone dysplasia. Cytogenetic studies showed a 46,XY,t(14;18)(q13;q12) karyotype in the proband. The parental karyotypes are normal.

Case 10 (Cell line L6504)

The proband was a white male, born at 41 weeks gestation to a 21-year-old gravida 1 mother. His birth weight was 8 lbs 3 oz and microcephaly was noted at birth. At 7 months his head circumference was 40.4 cm (50th percentile for 4 months) and length was 63.8 (40th percentile for 4 months). He had a flat face, slightly low set ears and bilateral epicanthal folds. His limbs were normal and he demonstrated no ocular anomalies at the age of 7 months. His tone was increased with arching when held and had feeding and weight gain problems. At 13 months, he could not yet sit or walk, had poor head control and was able to smile but not follow. His height was 50th percentile for 5 months and head circumference was 50th percentile for 4 months. At 19 months his length was 50th percentile for 11 months and head circumference was 50th percentile for 5 months. Furthermore, he did not exhibit any known dental anomalies at this age. Cranial imaging showed agenesis of the corpus callosum. The patient died during his early childhood. Both parents are clinically normal and have normal karyotypes. The proband's karyotype was ascertained as 46,XY,t(4;14)(q25;q13). No complex rearrangement, other than a translocation, was apparent from his karyotype.

Case 11 (Cell line SMOM)

This is a Caucasian female who was clinically normal except with mild mental retardation and schizophrenia. No cranial imaging was done. Her karyotype was 46,XX,t(9;14)(q34;q13).

Case 12 (Cell line L6874)

The proband was a Caucasian female diagnosed with severe mental delay and schizophreniform psychosis with hallucinations, disturbed agitated behavior, and emotional lability. No cranial imaging was done. Her karyotype was 46,XX,t(9;14)(q34;q13). Her mother (case 11) is described above. Her father, deceased, was diagnosed with bipolar disorder II. She also has a sister with the same karyotype and has severe mental delay. No cell line or DNA specimens were available from the other sib and the father.

Case 13 (DNA specimen D2502)

This case (Caucasian), previously reported (Devriendt et al., 1998a), was born to term after an uncomplicated pregnancy, with a weight of 3400 g, length of 48 cm, and OFC of 36 cm. She had respiratory distress with recurrent infections, and high serum thyrotropin concentration without overt hypothyroidism. At the age of 2 yrs, she showed mental and motor developmental delay, was hypotonic and had truncal ataxia. An MRI analysis showed no brain abnormalities. Her karyotype was 46,XX,del(14)(q13q21).

Case 14 (Cell line LMEHTADC)

The Caucasian proband was born to full term weighing 3860 g, after a caesarian section delivery. During the neonatal period, he had broncho-pulmonary dysplasia, gastro-esophageal reflux, and congenital hypothyroidism. His motor and speech development were mildly delayed and he had mild craniofacial signs which included frontal prominence, downslanting palpebral fissures, and high arched palate. He also had generalized hypotonia, dysplasia of the hip and, broad hands and feet. An MRI analysis of his neurocranium showed no abnormalities. His karyotype was 46,XX,del(14)(q13.2q21.2).

Case 15 (DNA specimen D2291)

This case, previously reported (Chen et al., 1997) was a chinese male terminated at 27 weeks of gestation with a weight of 1006 g, and length of 36.5 cm. He had microcephaly, cebocephaly, hypotelorism, low set ears, micrognathia, short neck and cryptorchidism. An autopsy showed alobar holoprosencephaly with arhinencephaly, agenesis of the corpus callosum, subdural hematoma of the cerebellum, bilateral clavicular fracture, hemothorax, hemopericardium, hemoperitoneum, left testicular agenesis and right adrenal hypoplasia. His karyotype was 46,XY,del(14)(q13q21.1). The parents were clinically normal and had normal karyotypes.

Case 16 (Cell line SADC)

The subject (Caucasian) was born to term by cesarean section with a weight of 3230 g (25th-50th percentile), length of 50 cm (50th percentile) and OFC of 33 cm (< 2nd percentile). His craniofacial signs were partly within the holoprosencephaly spectrum and included a head circumference which continues to grow 3 to 4 standard deviations below the normal. He also demonstrated brachycephaly, bilateral ptosis, strabismus, downward slanting palpebral fissures, high arched palate with lateral ridges and a prominent nasal bridge with narrow nostrils, which cause breathing difficulties. No obvious intracranial abnormalities were found on his cranial CT scan at the age of 2 months and later at 8 years, however the presence of minor abnormalities cannot be excluded. His developmental milestones appear normal with the exception of minor academic difficulties at school.

Case 17

The proband (Caucasian) was born to term by cesarean section. Her birth weight was 3140 g (25th -50th percentile), birth length was 50 cm (50th percentile) and OFC was 32 cm (< 2 percentile). Physical examination demonstrated craniofacial signs within the holoprosencephaly spectrum. These included microcephaly, hypotelorism, cleft palate and median cleft lip with agenesis of the philtrum, premaxillary alveolus and columella. Her developmental milestones appeared normal and she was healthy with the exception of feeding difficulty. She died suddenly at the age of eight months. An autopsy examination showed partial agenesis of the corpus callosum with associated malformed cingulate gyri and white matter heterotopia.

2.4. RESULTS

2.4.1. Mapping of aberrant chromosome breakpoints

Fifteen patients were cytogenetically characterized as having a deletion or translocation breakpoint involving the human chromosome 14q11.2-q22 region. The results are shown in Table 2-1. The deletion breakpoints for cases 1, 2, 5 and 13 were previously defined with low resolution by microsatellite typing (Shapira et al., 1994; Schuffenhauer et al., , 1999; Devriendt et al., 1998a). However in this study, the breakpoints were remapped more precisely using flow sorted chromosomes or a more dense panel of microsatellite markers. The deletion breakpoint junctions of cases 1 to 6 were determined by flow sorted chromosome analysis (with a resolution of < 0.5 to 1.5 Mb), while for cases 13 to 15, microsatellite typing was done (Table 2-1). For these cases, the resolution of the deletion breakpoint interval by use of flow sorted chromosomes far exceeded that of microsatellite typing, due to the lack a dense panel of microsatellite markers and the non-informativeness of these markers. For case 13, a more dense panel of microsatellite markers was used to genotype this patient in order to confirm the previous mapping by Devriendt et al. (1998a), and to determine whether the deletion interval could be more precisely defined. Subsequently, the mapping results were confirmed, however, a more precise definition of the deletion interval centromeric to D14S1014 was not possible. Examples of the microsatellite typings are shown in Figure 2-1a. For flow sorted chromosome analysis, the breakpoint junctions for deletion chromosomes are defined by the change in absence to presence, or vice versa, of a marker (Figure 2-1b),

Table 2-1: Patient mapping information, listing markers in 14q1-q2

Marker	case 1	case 2	case 3	case 4	case 5	case 6	case 7	case 8
Cell line/ DNA	L251DC	L777DC	L640DC	HQ0004	SAN	MAA	L6873	B01060
D14S1070	+ (71)	+ (67)	+ (100)	+ (100)	+ (100)			
D14S590			+ (100)					
D14S1220			+ (100)					
D14S64	+ (100)	+ (67)	+ (100)	+ (67)	+ (100)			
D14S275	+ (43)	+ (67)	- (100)	+ (67)	+ (100)			
D14S80		+ (67)	- (100)	+ (33)	- (67)			
D14S835				+ (100)		+ (100)		
D14S740				+ (100)				
D14S262		+ (67)	- (100)	- (100)	- (67)	+ (67)		
D14S975			- (100)	- (100)	- (100)	- (100)		
D14S1021	+ (100)	+ (67)	+ (67)	- (50)	- (100)	- (67)		
D14S297	+ (100)	- (67)	+ (67)	- (83)	- (100)	- (50)	+D14 (100)	
D14S1236	- (100)	- (100)						
D14S730	- (100)	- (83)	+ (100)	- (100)	- (100)	- (77)	+D14 (67)	
D14S49								
D14S70	- (100)		+ (100)	- (100)	- (100)	- (83)		+D14 (100)
D14S599						- (100)		
D14S988								
D14S1014				- (100)	- (100)	- (83)	+D14 (100)	+D14 (100)
D14S888				- (83)	- (100)	- (100)		+D14 (100)
D14S253			+ (33)	- (100)		- (100)		
D14S69					- (100)	- (67)		+D14 (100)
D14S75		- (100)		- (100)	- (100)	- (77)		+D14 (100)
AFM205XG5					+ (100)	- (100)		+D4 (100)
AFM200ZH4			+ (67)		+ (100)	- (100)		+D4 (100)
D14S306		- (100)		- (100)	+ (100)	+ (67)		+D4 (33)
D14S728				+ (100)	+ (100)			+D4 (100)
D14S286								
D14S738				+ (100)		+ (77)		
D14S278								
D14S600	- (67)	- (100)		+ (100)		+ (50)	+D14 (100)	
D14S872								
D14S777	- (75)							
D14S1013	- (100)	- (100)			+ (67)	+ (83)	+D14 (100)	
D14S266	+ (83)		+ (100)			+ (33)		
D14S288	+ (83)	- (100)				+ (100)		
D14S976	+ (67)	- (100)					+D14 (100)	+D4 (100)
D14S269	+ (67)	- (100)				+ (100)	+D4 (100)	
D14S978	+ (33)	- (100)				+ (83)	+D4 (100)	
D14S276	+ (29)	- (100)				+ (67)		
D14S52	+ (100)							
D14S66	+ (100)	- (100)		+ (100)				
D14S575		- (100)						
D14S592		+ (100)						
D14S1059		+ (100)						
D14S63		+ (67)						
D14S251	+ (100)					+ (67)	+D4 (100)	

Of note, the highlighted region represents the defined HPE8 locus.

+ = presence of marker tested
- = absence of marker tested
D = derivative chromosome
Ni = non informative

- mapping was done on flow sorted chromosomes for cases 1 to 8 three to nine times
- numbers in parentheses refer to percent of tests with final result
- the order of markers is from centromere (top) to telomere (bottom)

Table 2-1: Patient mapping information continued, listing markers in 14q1-q2

Marker	**case 9**	**case 10**	**case 11**	**case 12**	**case 13**	**case 14**	**case 15**
Cell line/ DNA	BV0416	L6504	SMOM	L6874	D2502	LMEHT ADC	D2291
D14S1070						+(100)	
14S590							
D14S1220							
D14S64							
D14S275							Ni (100)
D14S80							+ (100)
D14S835				+D14 (100)			
D14S740				+D14 (100)			+ (100)
D14S262						+ (100)	
D14S975				+D14 (100)			
D14S1021			+D14 (100)	+D14 (100)	+ (100)		
D14S297				+D14 (100)	Ni (100)		+ (100)
D14S1236	+D14 (100)						
D14S730			+D14 (100)	+D14 (100)			
D14S49			+D14 (100)	- (100)		+ (100)	+ (100)
D14S70	+D14 (100)		+D9 (100)	+D9 (71)	+ (100)	Ni (100)	- (100)
D14S599				+D14 (100)	+ (100)		
D14S988				+D14 (100)	+ (100)		Ni (100)
D14S1014	+D14 (67)		+D9 (100)	- (100)	- (100)	+ (100)	
D14S888	+D14 (100)		+D9 (100)	+D9 (100)		Ni (100)	
D14S253					- (100)	+ (100)	
D14S69	+D14 (67)	+D14 (100)		+D9 (100)		- (100)	- (100)
D14S75	+D18 (100)	+D14 (100)		+D9 (100)	- (100)	Ni (100)	
AFM205XG5	+D18 (100)	+D14 (100)					
AFM200ZH4		+D14 (100)					
D14S306	+D18 (100)	+D4 (100)					
D14S728	+D18 (67)	+D4 (100)					
D14S286		+D4 (100)					
D14S738		+D4 (100)					
D14S278		+D4 (100)					
D14S600			+D9 (100)				
D14S872		+D4 (100)					
D14S777		+D4 (100)					
D14S1013		+D4 (100)	+D9 (100)				
D14S266		+D4 (100)				+ (100)	+ (100)
D14S288					- (100)	+ (100)	
D14S976	+D18 (100)	+D4 (100)					+ (100)
D14S269							
D14S978		+D4 (100)					Ni (100)
D14S276							
D14S52							
D14S66							
D14S575							
D14S592							
D14S1059							
D14S63						+ (100)	Ni (100)
D14S251							

Of note, the blue highlighted region represents the defined HPE8 locus.
+ = presence of marker tested
- = absence of marker tested
D = derivative chromosome
Ni = non informative

- mapping was done on flow sorted chromosomes for cases 9 to 12 three to nine times
- mapping was done by microsatellite typing on cases 13 to 15, up to two times
- numbers in parentheses refer to percent of tests with final result
- the order of markers is from centromere (top) to telomere (bottom)

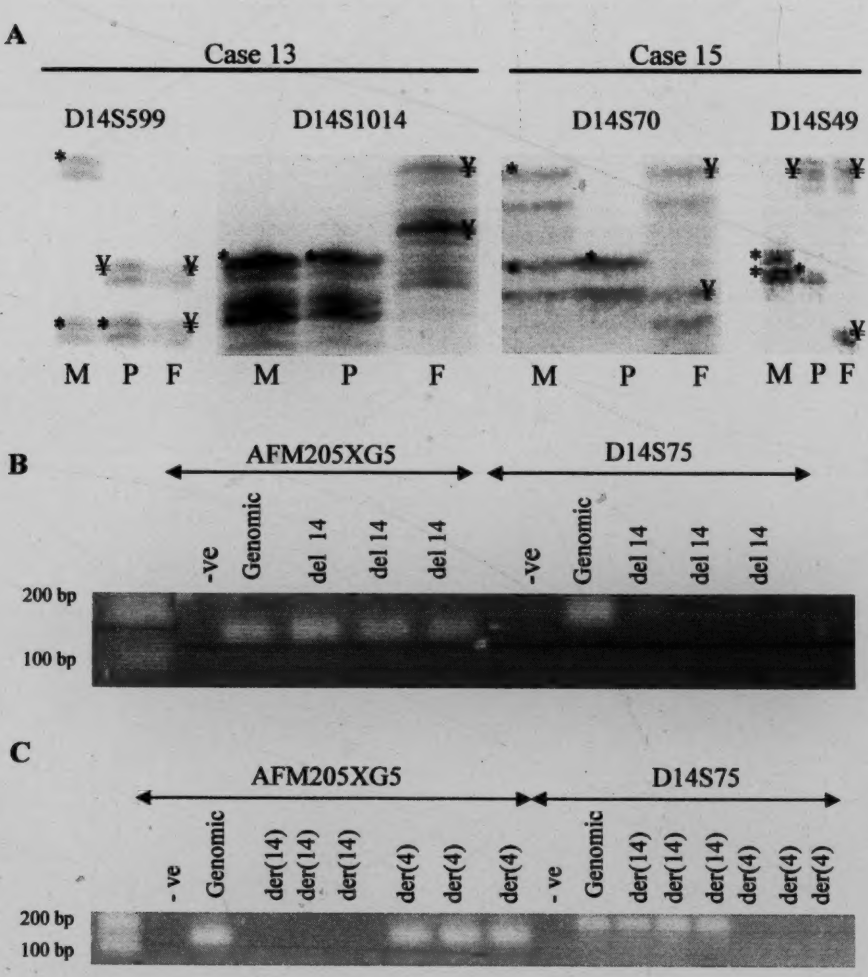

A
Case 13
Case 15
D14S599
D14S1014
D14S70
D14S49
M P F M P F M P F M P F
B
AFM205XG5
D14S75
-ve
Genomic
del 14
del 14
del 14
-ve
Genomic
del 14
del 14
del 14
200 bp
100 bp
C
AFM205XG5
D14S75
- ve
Genomic
der(14)
der(14)
der(14)
der(4)
der(4)
der(4)
- ve
Genomic
der(14)
der(14)
der(14)
der(4)
der(4)
der(4)
200 bp
100 bp

Figure 2-1: Examples of microsatellite typing and flow sorted chromosome analysis on selected patients. A) microsatellite typing on cases 13 and 15 showing paternal deletions of D14S70 and D14S1014. M=mother, P=proband, F=father, * = maternal allele(s), ¥ = paternal allele(s). B) flow sorted chromosome analysis of case five showing deletion of D14S75. del 14 = proximal 14 q deletion chromosome C) flow sorted chromosome analysis of case eight showing a breakpoint junction between AFM205XG5 and D14S75. der = derivative flow sorted chromosome.

while the breakpoint junctions of translocation chromosomes are defined by the switch of marker content from one derivative chromosome to the other (Figure 2-1c).

2.4.2. Defining the HPE8 minimal critical region

Six of the nine deletion patients used for mapping were affected with features within the holoprosencephaly spectrum. One had alobar HPE and the remaining five had features within the mild holoprosencephaly spectrum. It is noteworthy that the deletion interval of the patient with alobar HPE (case 15) overlapped with those manifesting features of mild holoprosencephaly (cases 1,2,4,5,6). The HPE8 minimal critical region was initially defined between the markers *AFM205XG5* and *D14S1014*, an estimated 4.2 Mb interval using the overlapping deletion intervals of a patient with alobar HPE (case 15) and a patient with features of mild HPE (case 5). By using the deletion intervals of patients not demonstrating HPE, that is, cases 3,13 and14, intervals not associated with holoprosencephaly can be excluded. Using this strategy, the deletion intervals of cases 13 and 14 were able to exclude the interval between the markers *D14S1014* and *AFM205XG5* thereby redefining the HPE8 minimal critical region between the markers *D14S49* and *D14S1014*, an estimated 2 Mb interval (Figure 2-2). Other recent molecularly defined deletion cases (Chen et al., 1998; Breedveld et al., 2002; Das et al., 2002; see also Table 2-2) also contribute to excluding proximal 14q regions not associated with HPE. Table 2-2 shows other constitutional proximal chromosome 14q deletion and translocation cases. One of the translocation and seven of the deletion cases have features within the holoprosencephaly spectrum. The molecular definitions of these cases are unknown.

None of the mapped translocation cases had features within HPE except case 10.

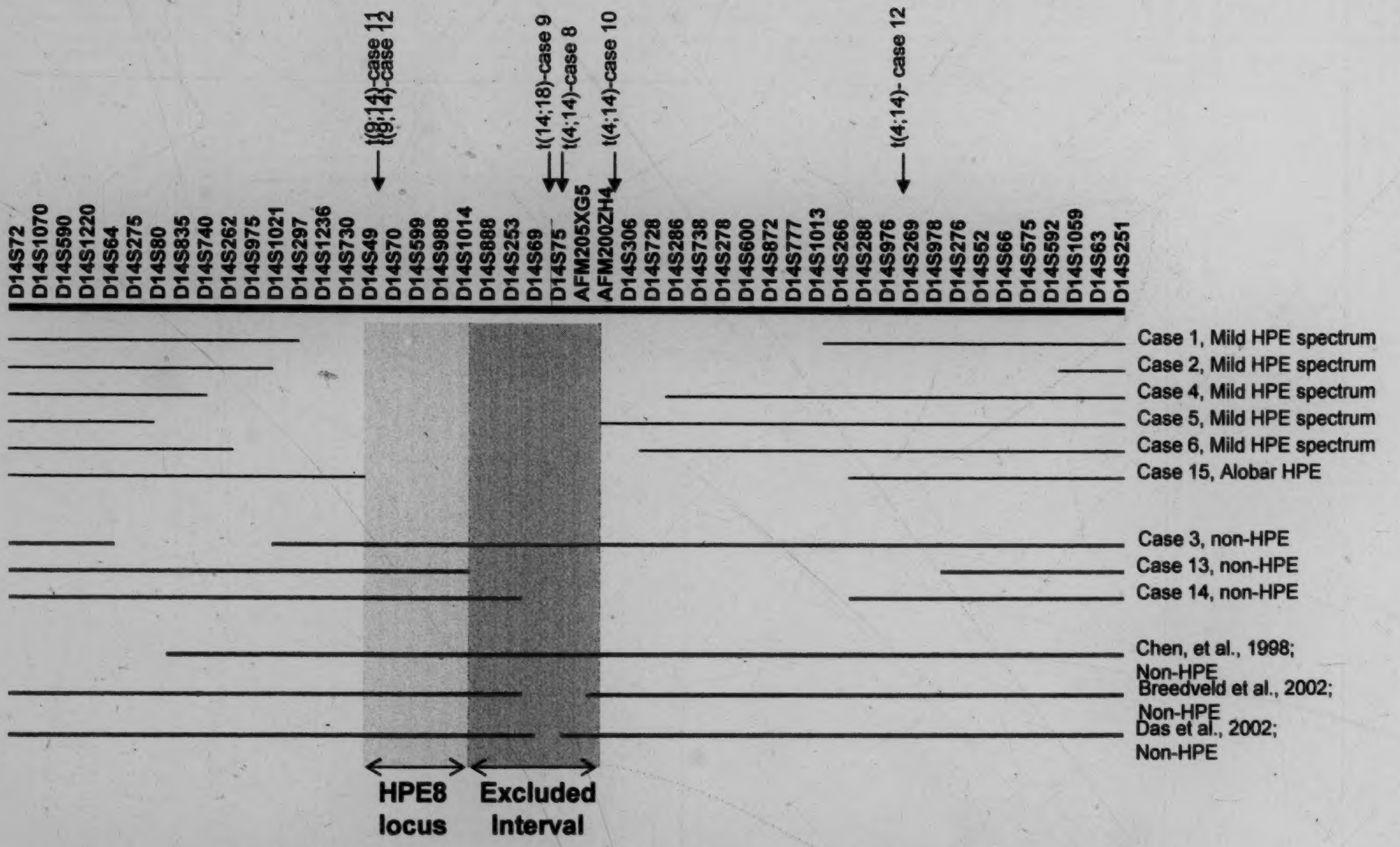

t(9;14)-case 11
t(9;14)-case 12
t(14;18)-case 9
t(4;14)-case 8
t(4;14)-case 10
t(4;14)- case 12
D14S72
D14S1070
D14S590
D14S1220
D14S64
D14S275
D14S80
D14S835
D14S740
D14S262
D14S975
D14S1021
D14S297
D14S1236
D14S730
D14S49
D14S70
D14S599
D14S988
D14S1014
D14S888
D14S253
D14S69
D14S75
AFM205XG5
AFM200ZH4
D14S306
D14S728
D14S286
D14S738
D14S278
D14S600
D14S872
D14S777
D14S1013
D14S266
D14S288
D14S976
D14S269
D14S978
D14S276
D14S52
D14S66
D14S575
D14S592
D14S1059
D14S63
D14S251
Case 1, Mild HPE spectrum
Case 2, Mild HPE spectrum
Case 4, Mild HPE spectrum
Case 5, Mild HPE spectrum
Case 6, Mild HPE spectrum
Case 15, Alobar HPE
Case 3, non-HPE
Case 13, non-HPE
Case 14, non-HPE
Chen, et al., 1998; Non-HPE
Breedveld et al., 2002; Non-HPE
Das et al., 2002; Non-HPE
HPE8 locus
Excluded Interval

Figure 2-2: Definition of the HPE8 locus. Black line=HPE spectrum cases; Blue line=non-HPE cases; Red line=characterized literature cases; yellow box=defined HPE8 locus based on overlapping deleted interval of all HPE phenotype spectrum cases mapped; green box=excluded HPE interval based on overlapping deleted regions of non-HPE cases; arrows point to translocation breakpoint junctions

Table 2-2: Other proximal human chromosome 14q deletion and translocation cases

Constitutional Cases	Clinical defect (minimal ascertainment)	Reference
46,X?,del(14)(q11.2q13.3)	Thyroid gland hypoplasia, severe respiratory distress, severe pulmonary distress, cystic mass besides pituitary gland	Krude et al., 2002
46,X?,del(14)(q13)	Benign hereditary chorea	Breedveld et al., 2002
46,XY,del(14)(q11.2q13.1)	Psychomotor delay, microcephaly, coarse facial features, spastic tetraparesis, agenesis of the corpus callosum	Ramelli et al., 2000
46,XY,del(14)(q11.2q13)	Microcephaly, mild dysmorphic face, short neck, arthrogryposis of the limbs, respiratory problems, mental delay, enlarged ventricles, white matter heterotopia, diabetes insipidus	Govaerts et al., 1996
46,XY,del(14)(q11q21)	Holoprosencephaly	Bruyere et al., 1996
45,XX,der(4)t(4;14)(p16.3q12)-14	Wolf-hirschhorn syndrome, hypertrophic cardiomyopathy, partial hemihypoplasia	Chen et al., 1998
46,XX,del(14)(q12q13.3)	Respiratory distress, developmental delay, joint contracture, hypertelorism, hypothyroidism, cerebral atrophy, ventricular dilatation, cyst of cavum septum pellucidum	Iwatani et al., 2000
46,XX,del(14)(q11.1q13)	Hypotelorism, premaxillary agenesis, semilobar holoprosencephaly	Levin and Surana, 1991
46,XY,del(14)(q11.1q13.2)	Holoprosencephaly	LeRoy, personal communication
46,X?,del(14)(q13)	Hypodontia	Das et al., 2002
46,XX,del(14)(q13;q22)	Cranial defect	Kamnasaran et al., 2001
46,XX,del(14)(q11.2;q13)	Agenesis of the corpus callosum, developmental delay, microcephaly	Kamnasaran et al., 2001
46,XY,t(8;14)(q22.3;q13)	Klippel-Trenaunay syndrome	Wang et al., 2001
46,XY,t(1;14)(p23.3q13)	Tetramelic mirror-image polydactyly, left inguinal hernia	Kim et al., 1997
46,XY,t(8;14)(p23.3;q13.3)	Developmental delay	Kamnasaran et al., 2001
46,XX,t(5;14)(q13;q12)	Hydrocephalus	Kamnasaran et al., 2001

Two of the six translocation cases used in the mapping had translocation breakpoint junctions within the HPE8 minimal critical region. Specifically, these were the familial t(9;14) cases 11 and 12 that had a translocation breakpoint junction close to the centromeric boundary of the minimal critical region between the markers D14S49 and *D14S70*. Case 12, the child of case 11, had additional complex rearrangements at 14q13 which included a deletion of the markers *D14S49* and *D14S1014*, which mapped about 2 Mb apart. It was not possible to determine if cases 11 and 12 have any forebrain anomalies.

2.4.3. Sex, ethnicity and parent of origin studies

Of the six patients with deletions who had features within the holoprosencephaly spectrum, the patient with alobar HPE (case 15) was a Chinese male. The remaining five patients with features of mild HPE (case 1,2,4,5,6) were predominantly females: one female was Hispanic (case 1), one female was Vietnamese (case 2), and three were Caucasians (two females and one male). The three other deletion cases with no HPE features (cases 3, 13, 14) were Caucasians (two males, one female).

The parental origins for the deletion chromosome of cases 1, 5, and 13 were previously defined by microsatellite typing (Shapira et al., 1994; Devriendt et al., 1998a; Schuffenhaeur et al., 1999). The parental origin for case 13 was confirmed in this present study. For case two, the parent of origin was ascertained from a polymorphic heteromorphism judged from GTG chromosome banding (Shapira et al., 1994). The parental origin for the deleted chromosomes of cases 3, 4, 6, 14 and 15 were determined in this study. All the deletions were de-novo cases. The parental origins for six of the nine deletion cases were paternal (cases 1,3,3,5,6,14 and 15), while cases 4 and 13

Table 2-3: Parent of origin analysis of cases mapped

Case	Type of aberration	Markers informative for deletion	Phenotype	Parental Origin of aberrant chromosome
1	Deletion	D14S54, D14S70	Mild HPE spectrum	Paternal
2	Deletion		Mild HPE spectrum	Paternal
3	Deletion	D14S975, D14S80	non-HPE	Paternal
4 *	Deletion		Mild HPE spectrum	Maternal
5	Deletion	D14S80, D14S70	Mild HPE spectrum	Paternal
6	Deletion	D14S70, D14S253	Mild HPE spectrum	Paternal
7	Translocation		non-HPE	?
8	Translocation		non-HPE	?
9	Translocation		non-HPE	?
10	Translocation		non-HPE	?
11	Translocation		?	?
12	Translocation		?	Maternal
13	Deletion	D14S1014	non-HPE	Maternal
14	Deletion	D14S69	non-HPE	Paternal
15	Deletion	D14S49, D14S70, D14S69	Alobar HPE	Paternal

* The undeleted markers, D14S275 and D14S66, were ascertained to be paternal in origin. See methods for determination of parental origin for this case. The parent of origin for the translocation cases were ascertained from cytogenetic studies on the families.

were maternal (Table 2-3). For those deletion cases exhibiting the holoprosencephaly phenotype spectrum, four were of paternal origin (case 1,2,5,6, and 15) and one was maternal (case 4). The parent of origin for one translocation case (case 12) was ascertained as maternal.

2.4.4. Analysis of family with a Robertsonian t(14q;22q) chromosome

A family with two sibs affected with microforms within the holoprosencephaly spectrum was ascertained by Drs. McLeod and Gerritsen (Alberta Children Hospital, Calgary). In this family (Figure 2-3a), the possibility of a locus, particularly on chromosome 14q13, as the cause of the holoprosencephaly spectrum of phenotypes was investigated. Lymphoblast cell lines from the parents and one of the affected child (case 16, Figure 2-3b) were used for molecular analysis. However, no DNA or cell line was available from the other affected child (case 17, Figure 2-3c) for molecular analysis.

2.4.4.1. *Karyotype analysis on case 16 and selected family members*

Standard cytogenetic analyses with G banding on the chromosomes of the family using cultured peripheral blood lymphocytes and/or prenatal amniocytes (collaboration with the Cytogenetics Facility (University of Alberta Hospital, Edmonton) showed the mother's karyotype was 45,XX,- 22,t(14;22)(q10;q10) (Figure 2-3d). The two affected sibs, a normal sib and a normal maternal uncle were also carriers of the Robertsonian (14q;22q) translocation chromosome, likely transmitted from one of the maternal grandparents (Figure 2-3a). All remaining family members have normal karyotypes. No deletions, inversions or complex chromosomal rearrangements were evident from karyotype analyses.

2.4.4.2. *Search for microdeletions*

Since Robertsonian translocations are frequent in the cause of additional rearrangements such as microdeletions, the chromosome 14q13 interval was analyzed for microdeletions in case 16. With a panel of 18 markers mapping to 14q13, analysis on flow sorted chromosomes demonstrated no deletions or complex rearrangements (Table 2-4a).

2.4.4.3. *Search for uniparental disomy*

Since Robertsonian translocation chromosomes are also frequent in the cause of uniparental disomies (UPD), chromosomes 14 and 22 were analyzed for UPD. With a panel of 17 microsatellite markers mapping to 14q1-q3, no evidence of maternal or paternal complete or partial uniparental disomy was found in case 16. Likewise, with a small panel of five microsatellite makers mapping to 22q11-q13, no evidence of maternal or paternal complete or partial UPD was found (Table 2-4b).

2.5. DISCUSSION

2.5.1. Mapping of chromosome rearrangements

Twelve chromosomal rearrangements involving the proximal region of human chromosome 14q have been defined within a physical distance of less than 0.5 to 1.5 Mb using flow sorted chromosomes. Six are deletion cases and the remaining six are translocation cases. The breakpoint intervals of another three deletion cases were mapped with a lower resolution using microsatellite typing. Based on the mapping results of these characterized deletion and translocation cases, no complex rearrangements were evident on chromosome 14. Complex rearrangements, such as the t(9;14) chromosome of case 12, were ascertained from a combination of positive and

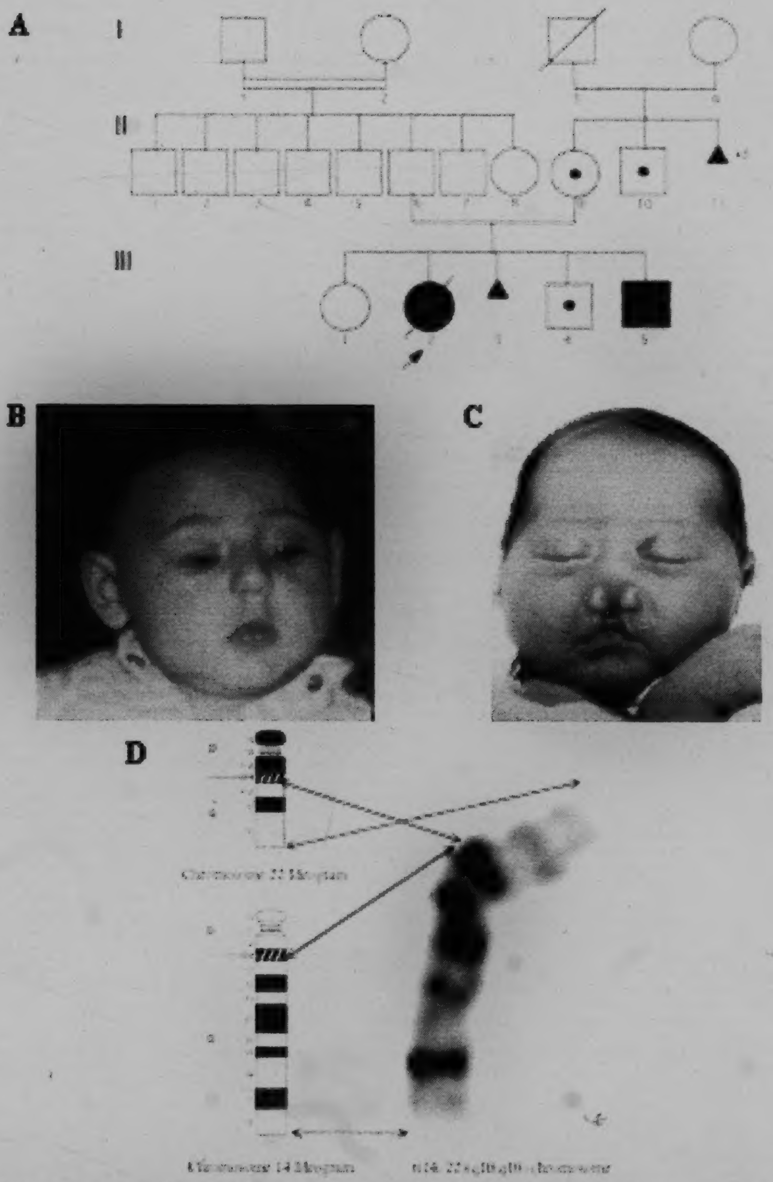


Figure 2-3: Family with Robertsonian translocation (14q22q) chromosome. A) pedigree of family with Robertsonian t(14q22q) chromosome. Case 16 is III-2; case 17 is III-5; B) case 16; C) case 17; D) partial karyotype of Robertsonian t(14q22q) chromosome at 400-500 GTG banding resolution. Permission was obtained by the parents to print the pictures of cases 16 and 17.

Table 2-4: Microdeletion and uniparental disomy analyses of case 16

a) Microdeletion analysis

Marker	**Results**	**Marker**	**Results**
D14S730	Not deleted (100)*	*D14S75*	Not deleted (100)
D14S49	Not deleted (100)	*AFM205XG5*	Not deleted (100)
D14S70	Not deleted (100)	*AFM200ZH4*	Not deleted (100)
D14S599	Not deleted (100)	*D14S306*	Not deleted (100)
D14S988	Not deleted (100)	*D14S728*	Not deleted (100)
D14S1014	Not deleted (100)	*D14S286*	Not deleted (100)
D14S888	Not deleted (100)	*D14S738*	Not deleted (100)
D14S253	Not deleted (100)	*D14S278*	Not deleted (100)
D14S69	Not deleted (100)	*D14S600*	Not deleted (100)

* number in parentheses refers to percent of test with final result

b) Uniparental disomy analysis

Marker	**Results**	**Marker**	**Results**
D14S49	Biparental inheritance	D14S48	Biparental inheritance
D14S70	Biparental inheritance	D14S81	Biparental inheritance
D14S599	Non-informative	D14S51	Biparental inheritance
D14S988	Non-informative	D14S45	Non-informative
D14S1014	Non-informative	D14S272	Biparental inheritance
D14S596	Non-informative	D14S1007	Non-informative
D14S306	Biparental inheritance	D22S431	Biparental inheritance
D14S69	Non-informative	D22S1265	Biparental inheritance
D14S52	Biparental inheritance	D22S1045	Non-informative
D14S63	Biparental inheritance	D22S417	Non-informative
D14S43	Biparental inheritance	D22S532	Non-informative

negative results for a set of markers and/or double positive or negative results for a set of markers on each derivative chromosome. For all the translocation cases, the possibility of a microdeletion at the translocation breakpoint junction on chromosome 14 cannot be excluded. Likewise, the possibility of complex rearrangements on the other unmapped derivative chromosome cannot be excluded. For the deletion chromosomes, the possibility of micro-rearrangements outside the cytogenetically characterized deleted region, regions not fully tested, cannot be excluded.

2.5.2. Parent of origin studies

Evidence for imprinting effects associated with human chromosome 14 is primarily supported by subjects who are carriers of a maternal or paternal uniparental disomy of chromosome 14, and are affected with a spectrum of phenotypes (reviewed by Kamnasaran and Cox, 2002). Maternal UPD 14 subjects frequently demonstrate short stature, precocious puberty, hypotonia, small hands, short philtrum, mental/motor delay, scoliosis, hydrocephalus and hyperextensible joints. Paternal UPD 14 subjects frequently have mental/developmental delay, small thorax, short neck, hairy forehead, blepharophimosis and polyhydramnios. Thus far, only the 14q32 interval has been shown to contain imprinted genes (review by Kamnasaran and Cox, 2002). Other intervals at the medial and proximal region of chromosome 14 are suggested to be imprinted, however further studies are warranted to confirm these notions (reviewed by Kamnasaran and Cox, 2002). In this study, the parent of origin for seven of the nine deletion cases examined is paternal, and two deletion cases are of maternal origin. Another three deletion cases of paternal origin are reported in the literature (Govearts et al., 1996; Ramelli et al., 2000). Interestingly, both cases have central nervous system and some facial anomalies in common with the mild HPE spectrum and six of the deletion cases examined (Cases 1,2,4,5,6, 15). Another three deletion cases of maternal origin were also reported in the literature (Chen et al., 1998; Iwantani et al., 2000). The central nervous system anomalies for two of these cases (Iwantani et al., 2000) were in common with the mild HPE spectrum and six of the deletion cases examined (Cases 1,2,4,5,6, 15). The parent of origin results in conjunction with the clinical findings obtained from the panel of deletion patients is insufficient to support the notion of imprinting for the proximal region of chromosome 14. Only low birth weight, short stature, hypotonia and mental/ developmental delay, features frequently associated with chromosome abnormalities, are in common with the UPD14 cases.

2.5.3. Properties of the HPE8 minimal critical region

In the present study, the deletion intervals of both alobar and mild HPE spectrum cases were found to overlap, suggesting a common cause in the molecular pathogenesis. The critical region for holoprosencephaly at chromosome 14q13, also known as HPE8, was defined between the microsatellite markers *D14S49* (centromeric boundary) and *AFM205XG5* (telomeric boundary), an estimated 4.2 Mb interval, using only subjects affected with the holoprosencephaly spectrum. However by the use of exclusion mapping, the minimal critical region was reduced to an estimated 2 Mb between the markers *D14S49* (centromeric boundary) and *D14S1014* (telomeric boundary). The *D14S1014* telomeric boundary was defined with a non-HPE case (case 13) and the *D14S49* centromeric boundary was defined with an alobar HPE case (case 15). Although the penetrance for HPE is estimated between 70 to 83% based on an autosomal dominant mode of inheritance in nonsyndromic HPE families (Cohen, 1989a,b; Odent et al., 1998), the *D14S49-D14S1014* interval is an appropriate region to commence searching for one or more holoprosencephaly genes on chromosome 14. Exclusion of the *D14S1014* and *AFM205XG5* interval is also supported by recently reported families with no holoprosencephaly and microdeletions in this interval (Breedveld et al.,2002; Das et al., 2002). Only two of the panel of six translocation cases, namely cases 11 and 12, have translocation breakpoint junctions within the HPE8 defined minimal critical region. Detailed characterization and discussion of these patients are described later in Chapter Five.

There is tremendous clinical variability of the HPE phenotypic spectrum seen among the panel of deletion patients with alobar HPE or the mild HPE spectrum. The clinical features of these characterized patients are most similar to those of the HPE5 locus (13q32) (Brown et al., 1995, 2001). The central nervous system anomalies varied from alobar HPE, partial or complete agenesis of the corpus callosum, underdeveloped pituitary gland, pachygyria, white matter heterotopia and delayed myelination. One patient (Case 15) had arhinencephaly, however for the remaining cases, the ability to smell is unknown and untested. Craniofacial findings included cebocephaly, hypotelorism, hypertelorism, hypertelorism, iris coloboma, cleft lip/palate, midface hypoplasia, low set/dysplastic ears, microcephaly, premature stenosis of the corona suture, and micrognathia. The extra-cranial anomalies included hypothyroidism, mildly dysmorphic limb digits, renal tubular acidosis, patent ductus arterosis, patent foramen ovale, hemopericardium, adrenal gland hypoplasia and testicular agenesis. The developmental milestones were similar to those of holoprosencephaly patients and included: poor weight gain, developmental delay, hypotonia, respiratory distress, sleep distress, feeding problems, seizures, improper eye contact and recurrent infections.

There were no significant correlations between the deletion sizes, sex or ethnicity and the severity of phenotype. Three of the affected cases were of Asian descent, and the remaining of Caucasian descent. Two of the cases were males and four were females. Epidemiological studies have reported a 3:1 ratio of affected females to males for alobar HPE, a 1:1 ratio of affected females to males for lobar HPE, and a preponderance of more severely affected females (Cohen, 1989a,b). The findings from the HPE spectrum cases reported in this chapter cannot be robustly compared with previously reported epidemiological findings due to the small sample size. No evidence for parent of origin bias was found for the HPE phenotypic spectrum at 14q13 since five paternal and one maternal deletion cases were found, thereby supporting the notion that the HPE8 locus was not imprinted. No parent of origin findings have been reported thus far for the other proposed HPE loci, except that the *SHH* gene of HPE3 was suggested to be imprinted (Suthers et al., 1999).

2.5.4. Family with Robertsonian t(14q;22q) chromosome

Subjects who are carriers of Robertsonian translocations involving chromosomes 14 and 22 are rare and usually are clinically normal. Any aberrant phenotypes associated with Robertsonian translocations are due to additional micro-rearrangements or uniparental disomy (Shaffer and Lupski, 2000). The molecular characterization of case 16 was an attempt to ascertain the cause of the phenotype which had microforms within the holoprosencephaly spectrum, that is, brachycephaly, microcephaly and midface hypoplasia. Specifically, case 16 was analyzed in search for a micro-rearrangement at

14q13. An analysis of the G banded t(14q;22q) chromosome showed no evidence of a deletion or complex rearrangement. However, it is noteworthy that the 400-500 banding karyotype resolution is too low to conclusively demonstrate evidence for subtle rearrangements. Using flow sorted chromosomes from this patient, no evidence for a deletion at 14q13 was found with 18 markers. No evidence for maternal or paternal UPD 14 or UPD 22 was found. This finding is supported by the fact that the clinical features of this child are inconsistent with the typical UPD 14 spectrum (Kamnasaran and Cox, 2002). In addition, no UPD 22 phenotype spectrum has yet been established. The deceased sib (case 17) of this patient had a more severe phenotype, with premaxillary agenesis and partial agenesis of the corpus callosum which were within the holoprosencephaly spectrum. Case 17 may have additional rearrangements not found in case 16. However, these possiblities cannot be resolved since DNA was unavailable from case 17. The only association of chromosome 22 with clinical features of the holoprosencephaly spectrum are from previously reported cases of chromosome 22 partial aneuploids (Back et al., 1980; Schinzel, 1981; Voiculescu et al., 1987) and DiGeorge-velocardiofacial syndromes at 22q11.2 (Wraith et al., 1985; Kraynack et al., 1999). The majority of the cranial and extra-cranial features found within these cases, however, differ from the anomalies found in the affected sibs. At this time, the genetic basis of the phenotype in this family remains unknown.

2.5.5. Phenotype correlations among selected deletion cases and tentative candidate genes

The sizes of the deletions as ascertained from the Marshfield sex-averaged linkage map, vary from 4 cM to over 27 cM physically representing about 4 Mb to over 27 Mb, if one cM is estimated to be equivalent to one Mb. The translocation patients used in this research share no common clinical features with the deletion patients examined, except for case 10. The majority of the mapped translocation breakpoint junctions therefore cannot be used to pinpoint candidate gene(s) for the affected phenotypes in the deletion cases. The molecular characterization of the aberrant chromosomes of cases 10, 11, and 12 are described in detail in Chapter Five.

The common clinical features of cases 4,5,6, 13, and 14 are thyroid dysfunction and respiratory distress. Four proximal chromosome 14q deletion cases (Krude et al., 2002; Govaerts et al., 1996; Iwatani et al., 2000) were also reported to show respiratory distress, thyroid dysfunction and as well as other findings. Hypothyroidism and respiratory distress are within the holoprosencephaly spectrum, however for these cases, *NKX2.1(TTF-1)* (thyroid transcription factor 1), located between *D14S70* and *D14S75*, and/or *FOXA1(HNF3A)* (hepatocyte nuclear activating factor 3 alpha), located between *AFM200ZH4* and *D14S306*, are excellent candidate genes for these phenotypes. This stems from the following evidence. A homozygous knock-out of the *Nkx2.1* gene of mice resulted in pups with agenesis of the thyroid gland and other defects (Kimura et al., 1996). A patient with congenital thyroglobulin defect and almost complete loss of *NKX2.1* expression was also reported (Perna et al., 1997). Finally, the *NKX2.1* gene is recently reported mutated in patients with benign hereditary chorea (Breedveld et al., 2002) and choreoathetosis, hypothyroidism and pulmonary problems (Krude et al., 2002). For the families affected with benign hereditary chorea, only one was diagnosed with congenital hypothyroidism and the others were untested (Breedveld et al., 2002). The second gene, *FOXA1(HNF3A)*, is implicated in the patterning of the foregut and lung during mouse embryogenesis (Overdier et al., 1994) and has been shown to activate the expression of *Nkx2.1* in lung epithelial cells *in vitro* (Ikeda et al. 1996).

For the majority of dominant disease genes found thus far at the proximal region of chromosome 14q (Kamnasaran and Cox, 2002; Kondoh et al., 2002), the associated phenotypes are not observed among the deletion cases analyzed here. For instance, oligodontia involving the primary or permanent molars is not observed among cases 1,2,4,5, 6, 13,14 and 15 which are haploinsufficient for the *PAX9* gene, but the probands may be too young to demonstrate such a phenotype. Only case 2 and two other reported cases (Iwantani et al., 2000) were noted to have reduced hearing loss, which may be a result of haploinsufficiency of the *COCH* gene. In general, the phenotypic heterogeneity demonstrated among the proximal chromosome 14q deletion cases can result from the different deleted genes, different genetic backgrounds, variable penetrance and the age of onset for the manifestation of the disease trait has not been reached.

2.6. SUMMARY OF FINDINGS

1) Patients with both alobar holoprosencephaly and features within the mild holoprosencephaly share a common deletion interval suggesting that there is a common molecular pathogenesis.

2) A holoprosencephaly minimal critical region was defined between *D14S49* and *D14S1014*, an estimated 2 Mb interval.

3) The clinical features of patients with the holoprosencephaly spectrum is tremendously variable.

4) There is no correlation between the size of deletion, sex, or ethnicity and the severity of the holoprosencephaly spectrum of phenotypes.

5) The HPE8 interval is unlikely to be imprinted.

6) No evidence for uniparental disomy or a microdeletion was found in a Robertsonian t(14q22q) patient with mild holoprosencephaly microforms.

CHAPTER THREE

CONSTRUCTION OF PHYSICAL AND TRANSCRIPT MAPS OF THE HOLOPROSENCEPHALY MINIMAL CRITICAL REGION

The following people have contributed to work in this chapter:

Mandeep Seckon and Sean Humphray (The Sanger Centre, UK) provided detailed assistance on the use of the FPC software and supplemental fingerprint information on selected BAC clones.

Seldon Briand (Canadian Bioinformatics Resource Centre, Canada) wrote Perl scripts for the FPC software and database.

Undergraduate students Jai Shah and Aaron Davidson assisted with sequencing and/or annotating a selected region of chromosome 14q13.

3.1. INTRODUCTION

In Chapter Two, the minimal critical region for the HPE8 locus was defined between the markers *D14S1014* and *D14S49* by aligning the mapped deletion intervals of patients with proximal chromosome 14q rearrangements. In this chapter, YAC and BAC physical contigs are constructed in the HPE8 locus and a transcript map of this region is determined. These findings were used to test the sub-hypothesis that *"the HPE8 locus contains one or more candidate genes for holoprosencephaly"*. The region between the markers *D14S1014* and *AFM205XG5* that was excluded based on patients not exhibiting the holoprosencephaly phenotype spectrum (in Chapter Two) was also reexamined for potential holoprosencephaly candidate genes. Together, the findings of this chapter have identified candidate transcripts for holoprosencephaly at the HPE8 locus.

3.1.1. The physical and transcript mapping of other holoprosencephaly loci

The first holoprosencephaly locus to be physically mapped was HPE3, at 7q36 (Belloni et al., 1996). The HPE3 locus was about 500 kb in size. A contig comprising YAC, PAC and cosmid clones flanked by the markers *D7S550* and *D7S3024* was constructed. In order to identify transcripts within this interval, cDNA selection and exon trapping were performed. The *SHH* gene was mapped to the contig as part of the large scale physical mapping of human genes in the Human Genome Project. Interestingly, none of the trapped exons or selected cDNAs were part of the *SHH* gene. The *SHH* gene was subsequently found to map within 250 and 15 kb of the breakpoints from two translocation patients with the holoprosencephaly spectrum (Belloni et al., 1996).

HPE4 at 18p11.3 was the second locus physically mapped. Somatic cell hybrids were generated from six deletion cases and used to define a minimal region less than six

Mb in size (Overhauser et al., 1995). No physical or transcript maps were constructed at this interval since it was too large. Subsequently, the *TGIF* gene was mapped to 18p11.3 by metaphase FISH, as a suspected candidate gene, and was later found in the smallest deleted interval of an additional two HPE patients (Gripp et al., 2000).

HPE1 at 21q22.3 was identified using somatic cell hybrids established from three deletion HPE cases (Muenke et al., 1995). The minimal critical region was defined between *D21S113* and 21qter, an estimated 4.5 to 4.9 Mb interval. The *SIM2* (single minded homolog 2) gene was excluded as a candidate, by mapping outside of the minimal critical region. The lanosterol synthetase gene was later mapped to 21q22.3 using a somatic cell hybrid panel (Young et al., 1996). This gene, involved in the cyclization step of cholesterol biosynthesis, was assessed as the cause for recessive holoprosencephaly at HPE1. Mutations affecting the function of the lanosterol synthetase protein were not discovered among a panel of 30 HPE patients from 20 unrelated families (Roessler et al., 1999).

HPE2 was the fourth locus physically mapped. Six deletion, and three translocation cases were used to define a critical region between *D2S119* and *D2S88/D2S391* (Schell et al., 1996). A 1 Mb YAC contig was subsequently constructed at this interval. YAC clone 957B10, about 1.18 Mb in size, spanned the translocation breakpoint junctions of two HPE translocation cases. Mouse-human comparative mapping had suggested that the *SIX3* gene was within the HPE2 critical region. Subsequent isolation of cDNAs from the EST databases and mapping of the *SIX3* containing cosmid clone showed it mapped in HPE2, and within 200 kb of HPE translocation patients (Wallis et al., 1999).

HPE5 at 13q32 was initially defined as part of the 13q⁻ deletion syndrome, between the markers *D13S136* and *D13S147* (Brown et al., 1995). However, this interval was questionable since patients with the same deletion had other major defects such as exencephaly, and digital anomalies. A YAC contig was constructed and used to estimate the size of the HPE5 locus within 1 Mb. Two deletion patients with classic HPE later defined the minimal critical region as 300 kb in size. Using mouse-human comparative mapping, the *ZIC2* gene was suggested to map to 13q32. This gene was later found deleted in 13q32 deletion cases with the HPE spectrum, and also mutated in HPE patients with normal chromosomes (Brown et al., 1998).

Other HPE loci are yet to be defined more precisely. Given that the human genome is being sequenced with a tiling path of BAC clones, no physical contigs will be need to be constructed for these loci. Transcript maps will be constructed using gene annotation methods and the *in silico* mapping of newly isolated cDNA sequences.

3.1.2. Physical contigs at 14q13

At least 590 YAC clones were localized to chromosome 14 and arranged into five contigs by the mapping of a landmark of over 350 chromosome 14 Sequence Tagged Site (STS) markers, by efforts of the Whitehead Institute for Genome Research (USA) (Hudson et al., 1995). Of these YAC clones, at least 170 clones were binned to 14q13 using anchor points from markers genetically mapped to chromosome 14 (The genome database). Apart from this large scale physical contig mapping effort, only a 1.2 Mb BAC-PAC contig was reported by Matsumoto et al., (1997), in an attempt to find the gene responsible for tetramelic mirror image polydactyly. The tiling path of BAC clones reported through the Human Genome Project (Genescope) is also a physical contig at 14q13 that was reported during the later stages of this dissertation.

3.1.3. Transcript maps at 14q13

UniGene clusters were mapped in 1996 showing 16,354 unique ESTs within a genome-wide framework of about 1100 Généthon genetic markers (Schuler et al., 1996). A total of 434 ESTs mapped to chromosome 14, of which about 45 were binned to 14q13 (Schuler et al., 1996). In 1998, more than 30,000 non-redundant EST clusters were mapped on the G3 and/or the GB4 radiation hybrid panels (Schuler et al., 1996). 1,047 ESTs mapped to chromosome 14, of which at least 120 ESTs were binned to 14q13 (Deloukas et al., 1998). The availability of human chromosome 14 genomic sequence has now permitted the construction of transcript maps by the *in silico* alignment of Unigene clusters and newly isolated cDNA sequences to the genomic sequence. With this method, at least 85 genes map within an 8 Mb region at 14q13 (NCBI, Ensembl).

3.2. MATERIALS AND METHODS

3.2.1. Construction and mapping on YAC physical contig

CEPH YAC clones (791C9, 964F6, 937E3, 950B8, 825G12, 957H10 and 856F12) were selected from the Whitehead Institute for Genome Research YAC mapping project (www-genome.wi.mit.edu) based on having one or more markers binned between *D14S49* and *D14S1014*, the holoprosencephaly minimal region (HPE8). CEPH YAC clones were obtained from Research Genetics (Huntsville, Alabama). Glycerol stocks of YAC clones were plated on AHC or YPD solid medium and incubated at 30°C for 2 to 3 days until colonies appeared. From each plate, 4 to 6 recombinant single colony isolates were selected, pooled and grown overnight in 5 mls of AHC or YPD liquid medium shaking at 30°C, until the mixed cultures were turbid. Each mixed YAC clone culture was centrifuged at 4000 rotations per minute for 5 mins at room temperature. The liquid medium was discarded. The pellet was suspended in 800 µl of ddH_2O and boiled at 100°C for 20 to 30 minutes on a heat block. 2 µl of each pellet was used for subsequent whole cell PCR for the mapping of markers, ESTs and genes at 14q13. Specifically, each 20 µl PCR reaction contained 1.5 mM $MgCl_2$, 1X PCR buffer (Perkin Elmer), 50 ng primer, 0.2 mM dNTPs and 0.2 units of AmpliTaq (Perkin Elmer). The PCR primer sequences, conditions and cycles used in the mapping are listed in Appendix 4.

3.2.2. Construction of in silico BAC contig using the FPC software

The sequences for genes and markers that mapped onto the HPE8 YAC contig were subjected to BLAST searches to identify any sequenced BAC clones from the Human Genome Project. The markers *D14S1014* mapped onto BAC clone R173D09, *D14S70* mapped onto BAC clone R422L13, and *D14S49* mapped onto C2008K18. These three clones were used as chromosome walking reference points for the *in silico* construction of a BAC contig at HPE8. The software FPC (Fingerprinted contig) was downloaded from the Sanger Centre for Genome Research (ftp://ftp.sanger.ac.uk/pub/fpc/) and installed on a network of Sun Enterprise™ 4000 servers which was powered by the Solaris™ (SPARC™ Platform Edition) operating system. A database of over 300,000 fingerprinted BAC clones from the RPCI 1 and Caltech A, B, C and D libraries was also downloaded from Sanger Centre for Genome Research (ftp://ftp.sanger.ac.uk/pub/human/fingerprinting/hm/) and used with the FPC software. This database represented data from HindIII completely digested BAC clones. Using FPC, an *in silico* physical contig between the BAC clones R173D09 (distal) and C2008K18 (proximal) was constructed. Chromosome walking was done until overlapping clones were obtained with BAC 422L13 (a clone residing between R173D09 and C2008K18). Main clones are called parent clones. Clones that are of shorter sizes and shared at least five restricted fragments with a large clone are called sister clones. Chromosome walking was achieved by setting the following parameters with the FPC software: buried value from 0.1 to 0.05, tolerance value of 0.05 and varying the cutoff value between 10^{e-5} and 10^{e-24}. The buried value represents the percentage of HindIII digested bands that are unshared between sister clones and the parent clones. The tolerance value represents a window of "tolerance" necessary to determine whether two fingerprinted bands are the same. The cutoff value represents a value of identity between two clones based on fingerprint patterns. The fingerprint of a BAC clone was queried against the database of BAC clone fingerprints with the settings mentioned above. A list of possibly related BAC clones were obtained from such search using mathematical standards set by the FPC software and the fact that two clones are designated related if five or more restricted fragments are shared within a defined tolerance value (Soderlund et al., 1997).

3.2.3. Production of BAC library filters for screening

A library of 95 BAC clones was selected from the *in silico* tiling path of BAC clones (section 3.2.2). BAC clones were obtained from BAC-PAC resources (Roswell Park Memorial Institute, USA) and Research Genetics. See Appendix 5 for the list of BAC clones. Stab cultures of BAC clones were plated on LB plus 20 µg/ml chloramphenicol and incubated overnight at 37°C. Multiple gridded colony lift filters were made as follows: 96 well microtitre well plates were filled with 150 µl LB with 20 µg/ml chloramphenicol and 15 µl Hogness freezing medium (280 mM $K_2HP0_4.3H_2O$, 15 mM

$Na_3C_6H_5O_7.2H_2O$, 68 mM $(NH_4)_2SO_4$, 13 mM KH_2PO_4, 44% glycerol). BAC single colony isolates were picked with sterile pipette tips and transferred to the appropriate well. The microtitre plates were incubated overnight at 37°C. A gridded stamper was used to transfer the colonies to 80 x 120 mm Hybond Nylon N^+ filters. The filters were incubated overnight at 37°C. The filters were denatured in 0.5 M NaOH/1.5M NaCl by placing on soaked Whatmann 3MM papers for 4 mins, then steamed for 4 mins on a steam bath. Neutralization was done with 4 mins incubation in 1M Tris, pH 7.4/1.5 M NaCl. Digestions were done by half hour incubation at 37°C in 1M Tris pH 8.5, 50 mM EDTA, 100 mM NaCl, 1% SDS, 5% Proteinase K. Filters were baked at 65°C for 2 hrs then soaked in 2XSSC before used.

3.2.4. Production of probes and screening of BAC library filters

A selection of BAC clones to be used as probes for screening the library filters was grown in 100 mls of LB plus 20 μg/ml chloramphenicol overnight at 37°C, before BAC DNA was prepared with the Midi Plasmid DNA Extraction Kit (Qiagen) and dissolved in 400 μl TE, pH 8.5. A 16 μl aliquot was taken and sequenced with the IRD700/800 SP6 and T7 labeled primers using the Thermo Sequenase Fluorescent Labeled Primer Cycle Sequencing Kit with 7-deaza-dGTP (Amersham Pharmacia Biotech). The reactions were run on a Licor DNA sequencer Long reader 4200. The sequence was analyzed for repeats using BLAST searches against the Alu database of GenBank. Primers were designed with Primer Premiere 3 (www-genome.wi.mit.edu/cgi-bin/primer/primer3.cgi) using the sequences that lacked repeats. To make the template for the probes, each 20 μl PCR reaction contained 0.2mM dNTP, 2mM $MgCl_2$, 1X PCR buffer (Perkin Elmer), 50 ng of each primer, 2 μl of BAC DNA template and 0.2 units of AmpliTaq (Perkin Elmer).

A list of the primer pairs, PCR conditions and cycles are in Appendix 6. PCR products were electrophoresed on 1% agarose gel (Gibco) stained in ethidium bromide, and the bands were excised and purified using the Gel Extraction Kit (Qiagen). About 100 ng of probe was labeled using the Random prime labeling kit (Amersham Pharmacia Biotech) with 5 µl of ^{32}P-α-dCTP (10 mCi/ml). The probe was denatured at 100°C for 5 minutes in Cot-1 DNA mix (15 µg COT-1 DNA (Gibco), 5X SSC, 0.1% SDS) before added to filters prehybridized in 10 mls of Church and Gilbert buffer at 65°C. Hybridization was done overnight at 65°C. Post hybridization washes were done with 2XSSC, 0.1% SDS twice at room temperature, followed by 0.2XSSC, 0.1% SDS twice at 65°C. The blots were wrapped in Saran Wrap and exposed for an appropriate period with Biomax MS or X-OMAT AR X-ray films.

3.2.5. Annotation of genomic sequence

The scaffold of individual BAC clones in the interval between the markers D14S49 and D14S1014 was downloaded from Genbank (www.ncbi.nlm.nih.gov/entrez) and the Golden path chromosome 14 sequence (genome.ucsc.edu). The genomic sequences were fragmented into 25 kb segments using DOS command line scripts written for the CHOP software and subjected to BLAST searches against the human EST and non-redundant databases of Genbank (E=0.0001, word size=11, masking for low complexity and human repeats), and electronic PCR (www.ncbi.nlm.nih.gov/genome/sts/epcr.cgi). The sequences for all the hits to the query genomic sequence were retrieved from Entrez (www.ncbi.nlm.nih.gov) and subjected to another round of BLAST searches against the non-redundant and high throughput genomic sequence databases (organism= Homo sapiens, E=0.0001, word size=11 and masking for low complexity and human repeats), to reassess the alignment. Only those alignments with ≥ 97% identity were recorded and considered authentic sources of chromosome 14q13 sequence.

Sequences flanked by the markers D14S49 and D14S1014 from the Golden Path sequence, and two scaffolds of genomic sequence from Celera Genomics (celera.com) were downloaded and subjected to Genscan analysis (bioweb.pasteur.fr/seqanal/interfaces/genscan.html). A selection of predicted cDNAs encoding ≥ 80 amino acids were subjected to BLAST searches against the non-redundant DNA and protein, and human EST databases of Genbank. Only those predicted genes that were identical to known genes or EST clusters were recorded and used for further analyses. The gene and EST content from the above two annotation methods were compared to filter redundancies.

3.2.6. RACE analysis

5' and 3' RACE were performed on a Marathon-ready human fetal brain cDNA library (21-30 wks gestation, 10 pooled Caucasian male and female) (ClonTech) to determine the expression of potential genes and to obtain 5' and 3' extended cDNA sequences. Primers for 5' and 3' RACE reactions were designed from partial cDNA sequences retrieved from Genbank, with melting temperatures between 65°C and 72°C using Primer Premiere 3. Each 50 µl reaction contained components as specified by ClonTech. One fifth of the RACE reaction was diluted to 200 µl with ddH$_2$O before used as a template for hemi-nested RACE to increase the DNA yield or nested RACE to increase the specificity of the amplification reaction. Each 50 µl hemi-nested reaction contained 0.2 µM AP-2 primer (ClonTech), 0.2 µM gene specific primer, 1X Advantage Taq PCR buffer (ClonTech), 0.2 mM dNTPs, 5 µl RACE template, and 1X Advantage 2

Polymerase mix (ClonTech). Each 50 µl nested RACE reaction contained 0.2 µM AP-2 primer (ClonTech), 0.2 µM Gene specific nested primer, 1X Advantage Taq PCR buffer (ClonTech), 0.2 mM dNTPs, 5 µl RACE template, and 1X Advantage 2 Polymerase mix (ClonTech). Primers used for 5' and 3' RACE or nested RACE, and PCR cycles are listed in Appendix 7. The hemi-nested and nested RACE reaction products were electrophoresed on 1% agarose gels (Gibco) stained in ethidium bromide, excised, gel purified using the Gel Extraction kit (Qiagen), and eluted in 30 µl ddH_2O. 10 µl of the purified product was extended with Poly(A) by incubating at 72°C in a 30 µl reaction with 1X PCR Buffer (InVitrogen), 0.1 mM ATP, and 1 unit Taq polymerase (Sigma), to facilitate cloning. The extended products were ligated into pCR4-TOPO TA vectors (InVitrogen) and transformed into TOPO 10 F chemically competent cells (InVitrogen) as specified by InVitrogen. Four white colonies were picked from each plate and incubated in 3 mls of LB/ Kanamycin (50 µg/ml) overnight at 37 °C. Plasmids were isolated with GenElute™ Plasmid Miniprep Kit (Sigma) and eluted in 100 µl water. To check whether a RACE product was cloned, 20 µl PCR reactions were setup with 2 µl of plasmid template, 1X PCR buffer (Perkin Elmer), 1.5 mM $MgCl_2$, 0.2 mM dNTPs, 50 ng M13 forward primer (cgccagggttttcccagtcacg), and 50 ng M13 reverse primer (agcggataacaatttcacacagga). The PCR cycle was: denaturation for 3 minutes at 94°C, 5 cycles of 94°C- 30 secs, 60°C - 30 secs, and 72°C- varying from 30 secs to 2 minutes, followed by 25 cycles of 94°C - 30 secs, 53°C - 30 secs, and 72°C – varying from 30 secs to 2 minutes. A final extension of 72°C for 5 minutes was done. PCR products were examined on 1% agarose gels (Gibco) stained with ethidium bromide. A selection of the plasmids with cloned RACE products was sequenced with the IRD700/800 M13 labeled primers (Amersham Pharmacia Biotech) using the Thermo Sequenase Fluorescent Labeled Primer Cycle Sequencing Kit with 7-deaza-dGTP (Amersham Pharmacia Biotech). Reactions were run on a Licor DNA sequencer Long reader 4200. The sequences were subjected to BLAST searches against the non-redundant and high throughput genomic sequence databases of Genbank to check whether the correct RACE product was cloned.

3.2.7. RT-PCR analyses on human fetal brain Poly A RNA

RT-PCR on human fetal brain Poly A RNA was performed to determine the expression of potential genes. 0.5 µg of human fetal brain Poly A RNA (20-33 wks gestation, 10 pooled Caucasian male and female) (ClonTech) was digested in 1X DNAse I buffer (Gibco) and 2 units DNase I (Gibco) for 15 minutes at room temperature, then incubated with 1.25 mM EDTA (Gibco) for 10 minutes at 65°C to inactivate DNase I. 500 ng of OligoDT (Gibco) or 500 ng of random decamers (Ambion) and 0.5 mM dNTP mix (Ambion) were added to DNase I treated Poly A RNA, heated at 65°C for 5 minutes, chilled on ice, and 1X first strand buffer (Gibco), 80 units RNAseOUT (Gibco) and 0.01 M DTT (Promega) were added. The mixture was incubated at 55°C for 2 minutes before adding 200 units of SuperScript II (Gibco). A control reaction had no SuperScript II (Gibco) added. The mixture was incubated at 55°C for 1 hr to synthesize the first cDNA strand, heated at 70°C for 15 minutes to inactivate SuperScript II (Gibco), then diluted to 140 µl with ddH_2O. 2 µl of the RT products were used to setup standard 20 µl PCR reactions with primer pairs and conditions listed in Appendix 8.

3.3. RESULTS

3.3.1. Construction of a YAC contig

Seven CEPH YAC clones were selected from the Whitehead Institute for Genome Research (WIGR) database, having non-ambiguous, or ambiguous hits for markers binned within the *D14S49* to *D14S1014* interval. A YAC contig spanning about 3 Mb was made between the markers *D14S1236* and *D14S888* with six of these seven CEPH YAC clones (Figure 3-1). The HPE8 critical region resided within this interval. The sizes of the YAC clones, taken from WIGR, varied from 1.1 Mb to 1.5 Mb. The size of one YAC clone was undetermined. It is unknown whether the YAC clones have microdeletions or undetected chimerism. One YAC clone was chimeric since STS content mapping (WIGR) showed markers mapping to chromosome 14 and to other chromosomes. Of the panel of genes that was used for physical mapping at 14q13, only *PSMA6* (proteosome macropain subunit alpha 6) and *NFKBIA* (nuclear factor kappa B inhibitor of alpha) mapped onto the YAC contig. *PSMA6* was localized on YAC clones 937E3 and 957H10, while *NFKBIA* was localized on YAC clones 791C9 and 950B8, placing them between *D14S70* and *D14S888*. Eleven markers were mapped onto the YAC contig and placed into three bins (Figure 3-1).

3.3.2. Construction of a BAC contig

The markers and genes mapped onto the 3 Mb YAC contig served as a framework of anchor points to construct a BAC contig. The panel of mapped markers and genes from the YAC contig was subjected to BLAST searches to identify any possible genomic clones that were sequenced by the Human Genome Project. BAC clones R173D09 and C2008K18, randomly sequenced nucleation clones from Genescope, contained markers

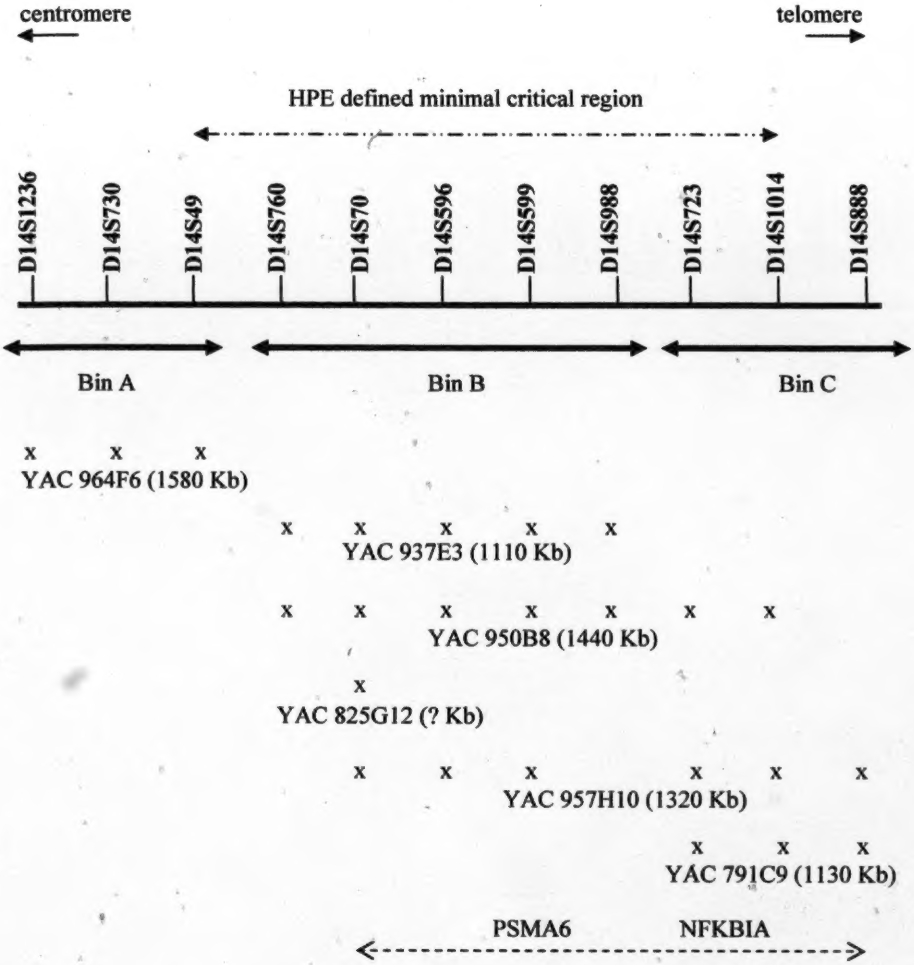

centromere
telomere
HPE defined minimal critical region
D14S1236
D14S730
D14S49
D14S760
D14S70
D14S596
D14S599
D14S988
D14S723
D14S1014
D14S888
Bin A
Bin B
Bin C
x x x
YAC 964F6 (1580 Kb)
x x x x x
YAC 937E3 (1110 Kb)
x x x x x x x
YAC 950B8 (1440 Kb)
x
YAC 825G12 (? Kb)
x x x x x x
YAC 957H10 (1320 Kb)
x x x
YAC 791C9 (1130 Kb)
PSMA6
NFKBIA

Figure 3-1: Three Mb YAC contig spanning the HPE8 minimal critical region. x=presence of a marker tested positive on YAC clones; the sizes of the YAC clones (in brackets) are taken from the Whitehead Institute for Genome Research YAC database; YAC clone 937E03 is chimeric; the locations of genes mapped onto specific YAC clones are demarcated by arrows; Bins A, B and C represent markers that mapped onto the same YAC clones. The order of markers in each bin is unknown.

D14S1014 and *D14S49* respectively. The marker *D14S70* was found on BAC clone *R422L12*, sequenced by WIGR. These three clones served as anchor points for the construction of a BAC contig using a two stage chromosome walking approach. Firstly, the *in silico* tiling path of over 300 fingerprinted BAC clones covering the interval from *D14S49* to *D14S1014* was made with the software FPC (Figure 3-2). Secondly, 95 BAC clones were randomly selected from the *in silico* tiling path to construct a BAC contig experimentally. Out of the 95 BAC clones selected, 30 were present in the tiling path between BAC clones R73E17, and C2008K18, after chromosome walking (Figure 3-3). Fingerprint information on each BAC clone is stated in Table 3-1. The BAC clone sizes ranged from 68 kb to 172 kb, average 127 kb, based on HindIII fingerprint distance estimates. A comparison of the BAC end sequences of selected clones with those of Genescope's tiling path by BLAST searches showed co-linearity in the order of the tiling path. A comparison of the two contigs further showed 1x to 11x coverage of a given interval from BAC clones R73E17 and C2008K18. BAC clone R85K15 was taken from Genescope to fill in the gap between BAC clones R173D09 and R73E17. The non-redundant physical distance from BAC clones R173D09 to C2008K18, that is the HPE8 locus, was estimated about 2 Mb in size.

3.3.3. Construction of a transcript map

With the availability of draft human genomic sequence from the Human Genome Project and Celera Genomics, a transcript map encompassing the interval *D14S49* to *D14S1014* was constructed by annotating 2 Mb of sequence for known genes, potential genes and ESTs using two methods. The first method attempted to identify all possible predicted genes using the GENSCAN algorithm. With the GENSCAN software, 65 potential genes

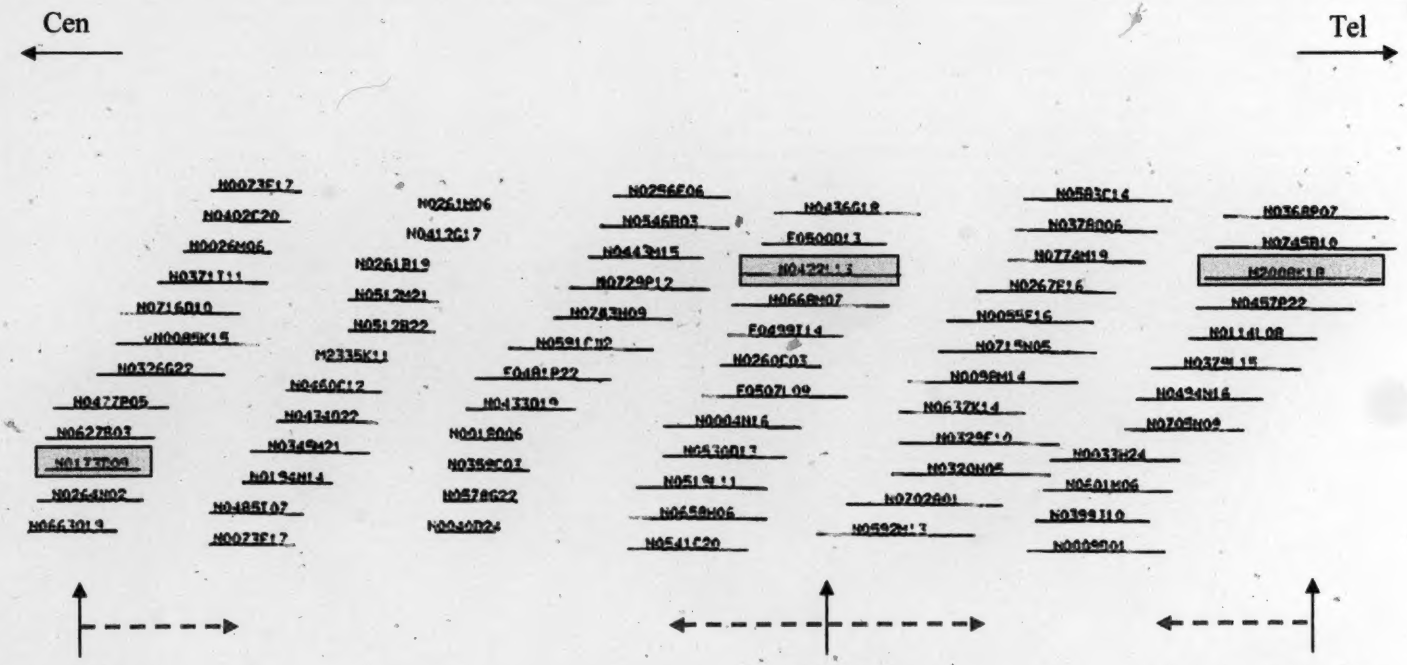

Cen
Tel
N0023E17
N0402C20
N0026M06
N0371I11
N0716D10
vN0085K15
N0326G22
N0477P05
N0627B03
N0173E09
N0264N02
N0663O19
N0261M06
N0412G17
N0261B19
N0512M21
N0512B22
M2335K11
N0460E12
N0434O22
N0349M21
N0194H14
N0485I07
N0023E17
N0296E06
N0546B03
N0443M15
N0729P12
N0783H09
N0591C02
F0481P22
N0433O19
N0018O06
N0358C03
N0578G22
N0040D24
N0436G18
F0500O13
N0422L13
N0668M07
F0499I14
N0260C03
F0507L09
N0004N16
N0530D17
N0519L11
N0658H06
N0541C20
N0702A01
N0592M13
N0583C14
N0378O06
N0774M19
N0267E16
N0055E16
N0715N05
N0098M14
N0637K14
N0329E10
N0320N05
N0033H24
N0601M06
N0399I10
N0009D01
N0368P07
N0745B10
M2008K18
N0457P22
N0114L08
N0379L15
N0494N16
N0705N09
D14S1014
D14S70
D14S49

Figure 3-2: Construction of an *in silico* BAC contig of the HPE8 minimal critical region using the FPC software. Only parent clones are displayed. N=RP11 BAC clones; M=Caltech D2 BAC clones; F=RP13 BAC clones. Red lines= direction of chromosome walking. The markers D14S1014, D14S70 and D14S49, and BAC clones highlighted in green boxes are the reference points for chromosome walking.

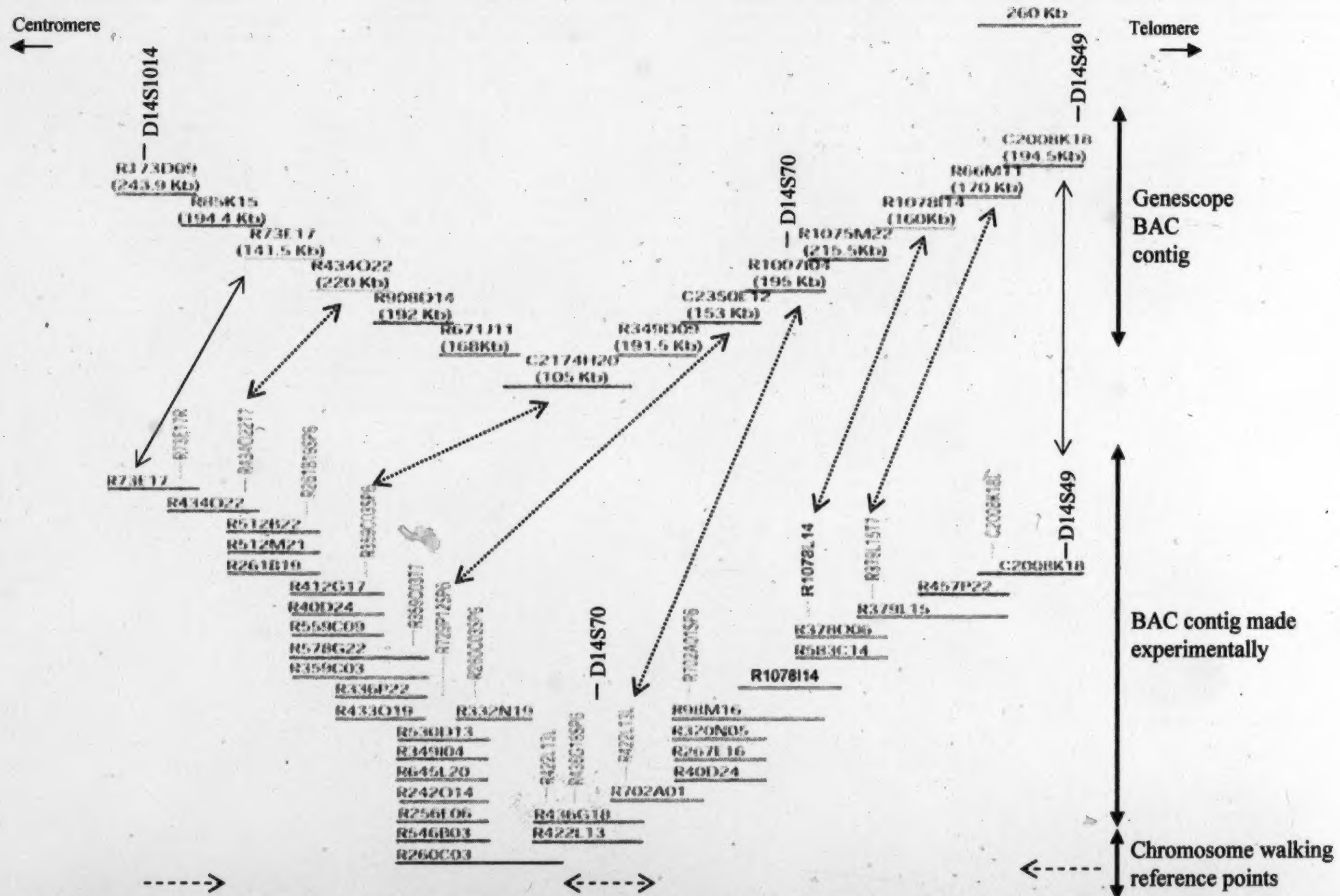

Centromere
Telomere
260 Kb
D14S1014
D14S70
D14S49
R173D09 (243.9 Kb)
R85K15 (194.4 Kb)
R73E17 (141.5 Kb)
R434O22 (220 Kb)
R908D14 (192 Kb)
R671J11 (168Kb)
C2174H20 (105 Kb)
R349D09 (191.5 Kb)
C2350E12 (153 Kb)
R1007I01 (195 Kb)
R1075M22 (215.5Kb)
R1078I14 (160Kb)
R66M11 (170 Kb)
C2008K18 (194.5Kb)
Genescope BAC contig
R73E17
R434O22
R512B22
R512M21
R261B19
R412G17
R40D24
R559C09
R578G22
R359C03
R336P22
R433O19
R332N19
R536D13
R349I04
R645L20
R242O14
R256L06
R546B03
R260C03
R436G18
R422L13
R702A01
R98M16
R320N05
R267L16
R40D24
R1078I14
R378O06
R583C14
R379L15
R457P22
C2008K18
BAC contig made experimentally
Chromosome walking reference points

Figure 3-3: Two Mb BAC contig of HPE8 minimal critical region. The markers D14S1014, D14S70 and D14S49 are the reference points for chromosome walking. Dashed arrows indicate directions for chromosome walking; Red arrows indicate anchor points between the built BAC contig and Genescope's BAC contig.

Table 3-1: Estimated sizes of BAC clones from the HPE8 physical contig tiling path.

BAC clone	# of HindIII digested fragments	Estimated size (Kb)
R434O22	28	112
R512B22	24	96
R512M21	25	100
R261B19	19	76
R412G17	12	48
R40D24	17	68
R559C09	38	152
R578G22	21	84
R359C03	22	88
R336P22	35	140
R433O19	30	120
R530D13	34	136
R349I04	33	132
R645L20	30	120
R242O14	31	128
R256E06	32	128
R546B03	35	140
R260C03	28	112
R332N19	33	132
R436G18	40	160
R422L13	44	176
R702A01	42	168
R98M16	40	160
R320N05	43	172
R267E16	38	152
R40D24	17	68
R378O06	39	156
R583C14	41	164
R379L15	41	164
R457P22	43	172
R1078I14	38	152

Estimated size= number of HindIII bands x 4000 bp, where 4000 bp is the frequency of a HindIII site. It is noteworthy that not all HindIII sites are functional. Therefore this formula calculates the minimum size of the BAC clone.

were predicted from the genomic sequence. 26 of the 65 predicted genes were found to be similar to known genes or EST clusters in the database. The second method attempted to identify all possible genes and EST clusters by performing BLAST and electronic-PCR searches on the genomic sequence. This method was used since the Genscan algorithm sometimes fails to predict genes. By using the second approach another eight possible genes were found. Hence, a total of 34 genes were found within the HPE8 locus (Table 3-2, Figure 3-4). Only 12 of the 34 genes are currently reported in the public annotation data of the human genomic sequence (Ensembl and National Centre for Bioinformatics Institute, July 2002).

Of the 34 genes, 13 were members of the reverse transcriptase family, likely due to mobile elements present within the 2 Mb region. Three genes (GS1RMCR, GS13RMCR, GS14RMCR) were similar to ribosomal proteins involved in translation. The *SRP54* gene (GS7RMCR) was known to be involved in the early stages of targeting nascent polypeptides to the rough endoplasmic reticulum for continued protein synthesis (Yin et al., 2001). One gene (GS5RMCR) was possibly involved in the function of the centromere based on the characterized yeast orthologue. The *BAZ1A* gene (GS4RMCR), is possibly a transcription factor involved in esophageal squamous cell carcinoma (Yasui et al., 2001) and shown to express in a variety of tissues such as spleen, thymus, prostate, skeletal muscle and testis, but not in adult brain (Jones et al., 2000). Two genes, namely *CFL2* (GS2RMCR) and *EGLN3* (GS24LMCR) were possibly involved in the structure, function and formation of the muscle (Thirion et al., 2001; Taylor, 2001).

All reverse transcriptase and ribosomal protein genes were excluded for further analyses as possible candidates for holoprosencephaly. The remaining 18 genes were

Table 3-2: Summary of genes found by annotating the 2 Mb of genomic sequence from D14S49 to D14S1014 (HPE8 candidate region)

Gene ID (2000, present study)	Identity	GenBank accession number (Protein/cDNA sequence)	Current (2002) NCBI/ Ensembl equivalent	Function
T75473	EST cluster	T75473		Unknown
*Gs31lmcr (NPAS3)	Neuronal PAS3	*AY157302, *AY157303	MOP6	Transcription factor with unknown function
Gs26lmcr	Leucine zipper motif-containing protein	AF181259.1		Possible transcription factor with neural linage role
Gs24lmcr	EGL nine homolog 3 (C. elegans)	NP_071356	EGLN3	Possible role in growth, differentiation and apoptosis in muscle cells.
H05452	EST cluster	H05452		Unknown
*R26150 (C14ORF11)	EST cluster	*AY157301	FLJ20578	Unknown
N99916 (SNX6)	Nexin sorting protein 6	NP_067072	SNX6	Trafficking of nodal receptor and nodal ser/thr receptor kinases
Gs4rmcr	Bromodomain adjacent to zinc finger domain, 1A	NP_038476	BAZ1A	Transcription factor with possible role in esophageal squamous cell carcinoma
AA609891	EST cluster	AA609891		Unknown
Gs2rmcr	Cofilin 2	NP_068733	CFL2	Muscle structural protein
Gs1rmcr	60S ribosomal protein L23a-like	P29316	LOC122585	Translation
Gs5rmcr	SMT3 suppressor of mif two 3 homolog 2-like	NP_008868		Possible centromeric protein
*Gs6rmcr (C14ORF19)	EST cluster with partial similarity to immunoglobulin binding proteins	*AY168620	LOC122587	Unknown
Gs7rmcr	Signal recognition particle 54kD	NP_003127	SRP54	Targeting of nascent translated proteins to rough endoplasmic reticulum
*Gs8rmcr, N67117 (C14ORF10)	EST cluster with similarity to calcium/calmodulin kinases	*AY157304, *AY157305	FLJ20644	Possible kinase
H41238	EST cluster	H41238		Unknown
Gs9rmcr	KIAA0391	NP_055487.1	KIAA0391	Unknown
Gs13rmcr	Ribosomal protein L7a-like	NP_038749.1		Translation
N50090	EST cluster	XP_101680.4	LOC161200	Unknown
N26711	EST cluster	N26711		Unknown
Gs14rmcr	Ribosomal protein L9-like	NP_000652.1		Translation
Gs11rmcr, Gs2lmcr, Gs4lmcr, Gs7lmcr, Gs12lmcr, Gs14lmcr, Gs15lmcr, Gs16lmcr, Gs25lmcr, Gs28lmcr, Gs30lmcr, Gs33lmcr, Gs44lmcr	Reverse transcriptase			Retroviral reverse transcriptase

Rows in Blue=not expressed in human fetal brain (21-30 weeks or 20-33 weeks gestation); Red=expressed in human fetal brain (21-30 weeks gestation), that is, candidates for holoprosencephaly; Green=not selected for further analyses based on possible function. Gene symbols in parentheses are official Human Gene Organization (HUGO) gene symbols. *= cDNA sequences isolated in this study and submitted to Genbank. All genes are listed from centromere (top) to telomere (bottom), except for the reverse transcriptases.

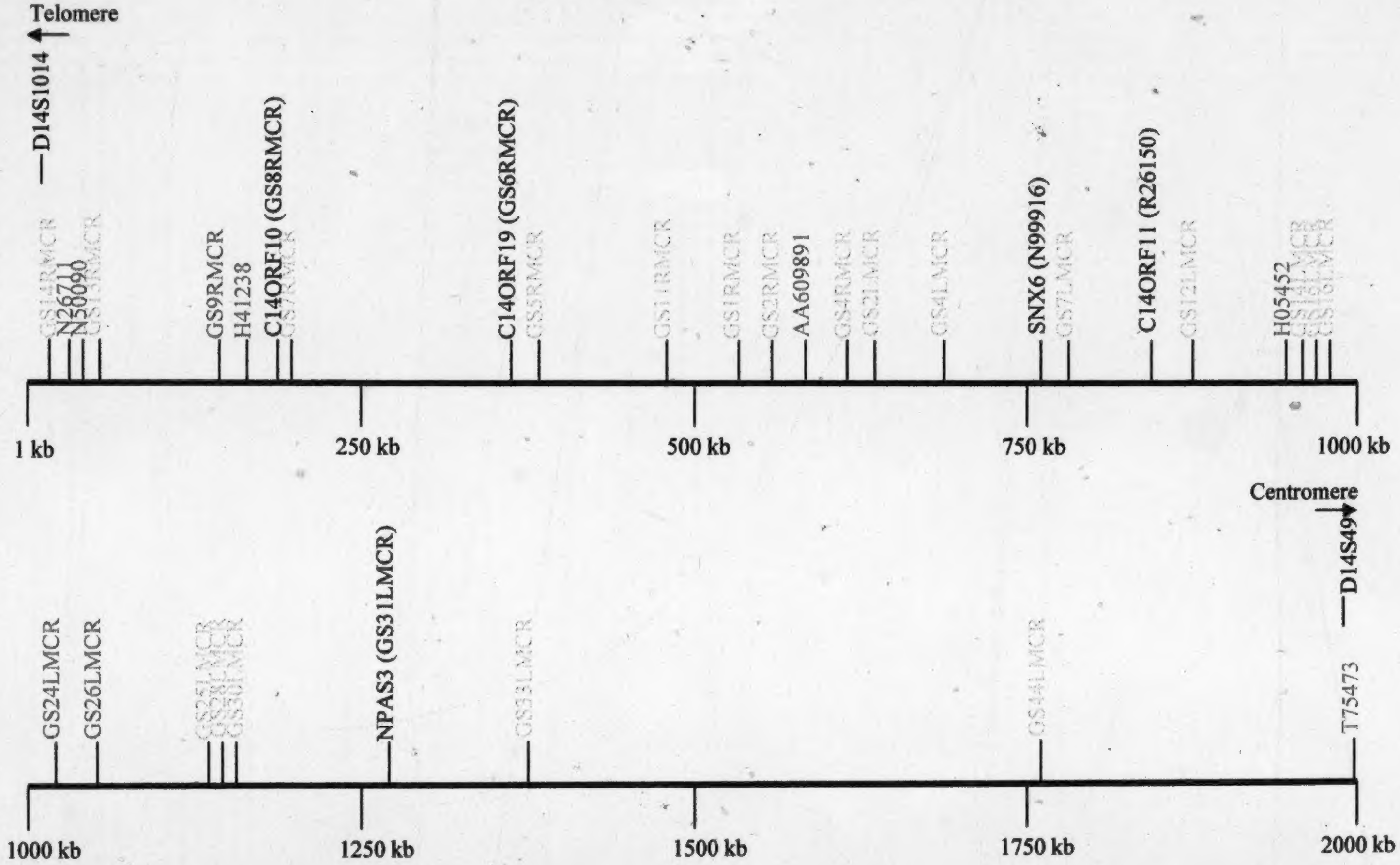

Telomere
D14S1014
GS14RMCR
N26711
N50090
GS13RMCR
GS9RMCR
H41238
C14ORF10 (GS8RMCR)
GS7RMCR
C14ORF19 (GS6RMCR)
GS5RMCR
GS11RMCR
GS1RMCR
GS2RMCR
AA609891
GS4RMCR
GS21LMCR
GS4LMCR
SNX6 (N99916)
GS7LMCR
C14ORF11 (R26150)
GS12LMCR
H05452
GS14LMCR
GS16LMCR
1 kb
250 kb
500 kb
750 kb
1000 kb
Centromere
D14S49
GS24LMCR
GS26LMCR
GS25LMCR
GS28LMCR
GS30LMCR
NPAS3 (GS31LMCR)
GS33LMCR
GS44LMCR
T75473
1000 kb
1250 kb
1500 kb
1750 kb
2000 kb

Figure 3-4: Transcript map of 34 genes annotated in the HPE8 locus. Only the 3' map positions of putative genes are shown. Genes shown in blue are not expressed in human fetal brain (21-30 weeks gestation); Genes shown in red are expressed in human fetal brain (21-30 weeks gestation), and are therefore candidates for holoprosencephaly; Genes shown in green were not selected for expression analyses.

selected for further analyzes as holoprosencephaly candidates by determining whether expression was present in human fetal brain (pooled 21-30 weeks gestation and/or pooled 20-33 weeks gestation). By using 5' and 3' RACE (hemi-nested and nested) on human fetal brain Marathon Ready cDNA library (21-30 weeks gestation), and RT-PCR on human fetal brain poly A RNA (20-33 weeks gestation), five of the 18 genes, namely *C14ORF19*, *C14ORF10*, *SNX6*, *C14ORF11* and *NPAS3*, were found to be expressed in human fetal brain. These five expressed genes were therefore candidates for holoprosencephaly at the HPE8 locus. One of the five genes, namely N99916, or Nexin sorting protein 6 (*SNX6*), has been characterized by the isolation of the cDNA sequence and by functional analyses of the protein (Parks et al., 2001; Ishibashi et al., 2001).

3.3.4. Examination of the *D14S1014* to *AFM205XG5* interval for potential holoprosencephaly candidates

The estimated 2.2 Mb region between the *D14S1014* and *AFM205XG5* markers was tentatively excluded in Chapter Two as a likely region for holoprosencephaly based on patients not exhibiting the holoprosencephaly spectrum and having deletions within this interval. The annotation efforts from NCBI and Ensembl (July, 2002) showed 12 genes resided in this interval. Two of the twelve genes, namely *NKX2.1* and *PAX9* were excluded for further analyses since they were mutated in hereditary benign chorea disorder and oligodontia respectively (Breedveld et al., 2002; Stockton et al., 2000). One of the remaining 10 genes was found expressed in human fetal brain (20-33 weeks gestation) by RT-PCR analysis (Table 3-3). This gene (FLJ23250 fis, GI: 10439870) was however an unlikely candidate for holoprosencephaly since it is similar to transposases.

Table 3-3: Expression analysis of genes mapping between *D14S1014* and *AFM205XG5**

Gene (NCBI MapViewer, Ensembl annotation, July 2002)	**Genbank accession number**	**Associated disease**	**Expressed in human fetal brain (20-33 weeks gestation)**
Proteosome macropain subunit alpha 6 (*PSMA6*)	NM_002791		No
Nuclear factor of kappa light polypeptide gene enhancer in B-cells inhibitor, alpha (*NFKBIA*)	NM_020529		No
Insuloma associated protein 1A6 (*INSM2*)	NM_032594		No
FLJ12926 fis (similar to neurofilament triplet M protein)	AK022988		No
FLJ23050 fis (similar to transposases)	AK026703		Yes
Breast cancer metasis suppressor (*MGC11296*)	NM_032352		No
Muk binding inhibitory protein (*MBIP*)	NM_016586		No
Homeobox Nkx2.9 (*NKX2.9*)	NM_014360		No
Thyroid transcription factor 1 (*NKX2.1*)	NM_003317	Hereditary benign chorea, choreoathetosis, hypothyroidism and pulmonary problems (Breedveld et al., 2002; Krude et al., 2002)	Not tested
Hypothetical protein (*MGC15427*)	BC008699		No
Mitochondrial 2-oxodicarboxylate carrier (*SLC25A21*)	NM_030631		No
Paired box 9 (*PAX9*)	NM_006194	Oligodontia (Stockton et al., 2000)	Not tested

*** Note this interval was excluded based on deletion patients not demonstrating holoprosencephaly features**

3.4. DISCUSSION

3.4.1. Construction of physical contigs at the holoprosencephaly minimal critical region

A YAC contig about 3 Mb in size was constructed and spanned the *D14S1014* to *D14S49* interval of the HPE8 locus. Only two human genes, namely *PSMA6* and *NFKBIA* mapped onto the YAC contig. The function of *PSMA6* is unknown, however with its similarity to proteosomes from macrophages, it is possibly involved in immune function like the *NFKBIA* gene (Haskill et al., 1991). Both genes are therefore unlikely candidates for holoprosencephaly. 12 markers were mapped onto the YAC contig and placed into three bins. A comparison of the physical placement of these bins with the current physical map positions of these markers on the human chromosome 14 genomic sequence showed agreement. The exact ascertainment of markers within the bins was not possible due to the large sizes of the YAC clones and also to the fact that YAC clones may have chimerism and microdeletions.

Three markers mapped onto the YAC contig served as anchor points in a chromosome walking approach to construct a BAC contig, about 2 Mb covering the interval *D14S1014* to *D14S49*. The strategy of easily constructing a BAC contig within a short time period in this study demonstrated the power of using fingerprint information of clones, and bioinformatics. The subsequent comparison of end sequences from this BAC contig with the BAC tiling path later reported by Genescope showed co-linearity in the order of the two contigs. The fact that two contigs of the HPE8 locus were made independently, and agreed, implied that any genomic sequence obtained from the tiling path of BAC clones from Genescope will be reliable with respect to the placement. The physical BAC contig constructed here will therefore be important in accessing the genomic sequence for microrearrangements.

3.4.2. Construction of a transcript map of the holoprosencephaly minimal critical region

A transcript map of the holoprosencephaly minimal critical region, comprising 34 genes, was constructed using bioinformatic analyses to annotate 2 Mb of genomic sequence for genes and EST clusters. With the annotation methods used, only potential genes with cDNAs or EST clusters in the database were discerned. Moreover, more than one discerned gene could possibly constitute a single gene. A comparison of the annotated 2 Mb genomic sequence with those from the public sequencing efforts (NCBI and Ensembl, July 2002) showed an additional eight genes (excluding the reverse transcriptases). In this 2 Mb interval, nine genes were reported by NCBI (July, 2002) and ten genes were reported by Ensembl (July, 2002). A total of only 12 genes were reported by these public databases when compared with those from this research. Interestingly, the *PSMA6* and *NFKBIA* genes that mapped on YAC clones spanning the HPE8 critical region, were unidentified in the 2 Mb annotated sequence. This is likely due to the fact that the Ensembl database showed the placement of these two genes outside the characterized HPE8 minimal critical region by mapping distal to *D14S1014* and proximal to *D14S888*. Any discrepancies between the annotation of genes and EST clusters are likely due to changes in the genomic sequence quality. Phases one or two draft genomic sequence of the 2 Mb HPE8 minimal critical region were used for annotation. The presence of ambiguous nucleotides, deletions, inversions and incorrectly placed sequences influence tremendously the accuracy of the Genscan gene predictions and to some extent the BLAST and electronic PCR searches on the sequence. As the quality of the genomic sequence improves, a more accurate annotation of the HPE8 interval will be obtained. Interestingly, an examination of the mouse-human comparative maps (Mouse chromosome 12 committee report, www.informatics.jax.org), and mouse genomic sequence (Ensembl, December 2002) showed no possible murine HPE candidate genes that could map to the human HPE8 locus.

Sixteen of the 34 genes, reverse transcriptases or ribosomal proteins, were excluded from further analyses as holoprosencephaly candidates. Of the remaining 18 genes, five were found expressed in human fetal brain (21-30 wks gestation) based on 5' and 3' RACE analyses. The five genes, namely *C14ORF19* (chromosome 14 open reading frame 19), *C14ORF10* (chromosome 14 open reading frame 10), *SNX6* (nexin sorting protein 6), *C14ORF11* (chromosome 14 open reading frame 11) and *NPAS3* (Neuronal PAS3) were therefore excellent candidates for holoprosencephaly. Further characterization of these five genes will be described in Chapters Four and Five. It is noteworthy that the central nervous system and facial architecture of the developing human fetus is established by 7 to 8 gestational weeks. It is assumed that a gene involved in central nervous system will continue expression in perhaps restricted regions of the developing fetal central nervous system. Given this, during 21-30 weeks of gestation, a period when the three fissures of the central nervous system form, a gene for holoprosencephaly could still express in fetal brain. From 30 to 33 weeks of gestation, secondary and tertiary sulci form in the developing fetal brain. The interval between the markers *D14S1014* and *AFM205XG5* that was tentatively excluded for harboring a holoprosencephaly locus in Chapter Two, was also analyzed for potential holoprosencephaly candidate genes. Only one of 10 genes, FLJ23250 fis, was expressed in human fetal brain (20-33 weeks gestation). The FLJ23250 fis gene, similar to transposases, suggested it was an unlikely candidate for holoprosencephaly. The *NFKBIA* and *PSMA6* genes, also mapping to this excluded interval, demonstrated no expression in human fetal brain (20-33 weeks gestation), and were unlikely holoprosencephaly candidates. In short, the 2.2 Mb excluded interval for holoprosencephaly between the markers *D14S1014* and *AFM205XG5*, did not have any striking candidate genes for holoprosencephaly thus far.

3.5. SUMMARY OF FINDINGS

1) YAC and BAC contigs were constructed in the HPE8 interval.

2) A transcript map comprising of 34 genes was identified in a 2 Mb interval of HPE8.

3) Five candidate holoprosencephaly genes at HPE8 were identified.

4) No striking candidate holoprosencephaly genes were found in the excluded *D14S1014* to *AFM205XG5* interval.

CHAPTER FOUR

CHARACTERIZATION AND FUNCTIONAL ANALYSIS OF HOLOPROSENCEPHALY CANDIDATE GENES

The following has contributed to work in this chapter:

Dr. Peter Dickie (University of Alberta Transgenics Centre) provided C57BL/6J pregnant mice for harvesting embryos.

4.1. INTRODUCTION

In Chapter Three, the 2 Mb minimal critical region for the HPE8 locus was annotated for genes and EST clusters. After excluding genes based on their possible functions and expression findings in human fetal brain (pooled 21-30 weeks gestation, and pooled 20-33 weeks gestation), five genes remained candidates for holoprosencephaly, namely *C14ORF19*, *C14ORF10*, *NPAS3*, *C14ORF11* and *SNX6* based on expression in the developing human fetal central nervous system. One of the five candidate genes, *SNX6* was characterized by other groups (Parks et al., 2000; Ishibashi et al., 2001; Teasdale et al., 2001). This chapter describes the isolation of the cDNA sequences for the other four of these five genes. Furthermore, all these genes were characterized with respect to determination of the genomic structure, proposal of properties based on extensive bioinformatic analysis, cellular localization and adult expression patterns, in order to gain insight into gene function. To further examine the spatio-temporal roles of these genes during murine embryogenesis, expression studies in mouse by RT-PCR analyses and whole mount *in situ* hybridization were done. The proposed experiments of this chapter were designed to test the sub-hypothesis that "*the HPE8 candidate genes have similar properties to other known holoprosencephaly genes*". The findings of this chapter have further characterized and selected excellent candidate genes for holoprosencephaly at the HPE8 locus.

4.1.1. Previous proposed candidate genes for holoprosencephaly on chromosome 14

The criteria for holoprosencephaly candidate genes are stated in Chapter One. Thyroid transcription factor 1 (*TITF1/NKX2.1*) was a proposed candidate gene for holoprosencephaly at 14q13 (Roessler and Muenke, 1998; Devriendt et al., 1998b). This transcription factor functioned in thyroid, lung, pituitary and ventral brain development based on the murine embryonic expression patterns and knock-out phenotype. Several lines of evidence however suggested that *NKX2.1* is not an HPE gene. Firstly, a transgene insertion into the *Nkx2.1* gene of mice resulted in pups with partial septal defects of the diencephalon and absence of ventral midline structures such as the mammillary bodies, but no other brain malformations within the severe to mild HPE spectrum were evident (Bingle, 1997). Secondly, the expression of this gene was almost completely lost in a patient with congenital thyroglobulin defect and no HPE (Acebron et al., 1995). Thirdly, heterozygous mutations have been recently reported in the *NKX2.1* gene in patients with choreoathetosis, hypothyroidism, and pulmonary alterations (Krude et al., 2002) and benign chorea disorder (Breedveld et al., 2002). Fourthly, the *NKX2.1* gene physically maps 1.4 Mb distal to the *D14S1014* most telomeric marker of the HPE8 locus. Fifthly, screening the *NKX2.1* gene for mutations in HPE patients failed to find any mutations (Maximillan Muenke, personal communication).

The human brain factor-1 gene (*FOXG1B/BF-1*) was also suggested as a candidate for holoprosencephaly at 14q13 (Maximillan Muenke, personal communication). The protein is a member of the winged-helix transcription factor family that functions in the differentiation of progenitor cells and patterning of the telencephalon (Li et al., 1996). The *Bf-1/Foxg1b* gene was expressed in telencephalic-derived structures such as the hippocampus, olfactory bulbs, basal ganglia and cerebral cortex (Duo et al., 1999). Homozygous null alleles of the *Bf-1/Foxg1b* gene resulted in pups that died at birth and had cerebellar hypoplasia (Xuan et al., 1995, Duo et al., 1999). Moreover, the development of the ventral telencephalon was more severally affected than the dorsal telencephalon (Xuan et al., 1995) due to aberrant proliferation and differentiation of telencephalic neuroepithelial cells (Duo et al., 1999). The defects in dorso-ventral patterning of the telencephalon resulted in ectopic expression of *Bmp4* in the dorsal telencephalon neuroepithelium and loss of *Shh* expression in the ventral telencephalon (Duo et al., 1999). More thorough examinations of the homozygous *Bf-1/Foxg1b* knockout mouse phenotype showed lack of the developing optic vesicle and stalk which was replaced by an enlarged retina (Huh et al., 1999). Biochemical studies of human BF-1/FOXG1B protein suggested that it served as a transforming growth factor beta antagonist possibly by interacting with other SMAD2 interacting proteins such as FAST2, preventing SMAD2 from positively mediating the TGFß signaling (Dou et al., 2000). BF-1/FOXG1B may also recruit transcription factors such as Groucho and histone deacetylase to repress transcription (Yao et al., 2001). The *FOXG1B/BF-1* gene maps 4.6 Mb proximal to the D14S49 most centromeric marker of the HPE8 locus. Mutation analysis of the *FOXG1B/BF-1* gene in a panel of HPE patients was unsuccessful in finding any (Maximillan Muenke, personal communication).

The *BMP4* gene at 14q22-q23 was suggested as a holoprosencephaly candidate based on the fact that ectopic recombinant Bmp4 protein in the anterior neural tube of developing chick embryos resulted in the loss of the basal telencephalon, cyclopia and missing ventral midline structures (Golden et al., 1999). However, the *BMP4* gene maps to the medial cytogenetic region of chromosome 14, which shows no evidence of association with holoprosencephaly to date, based on chromosomal cases and genetically positioned diseases.

4.2. MATERIALS AND METHODS

4.2.1. Screening of λ TriplEx cDNA library

The λ TriplEx Human Fetal Brain 5' Stretch Plus cDNA Library (pooled 10 male/female, 20 to 25 wks gestation) (ClonTech) was screened. *Escherichia coli* XL1 Blue plating cells were prepared for transfection by growing at 37°C in LB/0.2% maltose/10 mM $MgSO_4$. The cells were diluted to OD=0.3/ml for transfection. 4×10^4 to 5×10^4 pfu were incubated with *E.coli* XL1 plating cells (OD=0.18) for 15 min at 37°C, before adding top agar (LB/agar/10mM $MgSO_4$), and plating on 150 mm LB/10mM $MgSO_4$ plate. The dried plates were incubated at 37°C for 10 hrs for plaques to form. About 1.9×10^6 pfu were plated in total. Plaque lifts were then performed on cooled plates as follows. Hybond N^+ filter (Amersham Pharmacia Biotech) was incubated on the plate for 2 mins and the orientation was marked with 3 asymmetric holes using a 22G 1½ needle. The membrane was removed and chemically treated with denaturation solution (0.5 M NaOH, 1.5 M NaCl) for 2 minutes, then in neutralizing solution (1.5 M NaCl, 0.5 M Tris-HCl) for 5 minutes. The plaque lift was rinsed briefly in 2XSSC, 0.2 M Tris-HCl and dried at 80°C for 2 hrs before used. The filters were hybridized with probes made from the following genes: *C14ORF19*, *C14ORF10*, *NPAS3* and *C14ORF11*. Probes to screen the cDNA library were made by standard PCR. PCR primers were designed from partial cDNA sequences obtained from database searches against Genbank and from 5' and 3' RACE, using Primer Premiere 3. The PCR primers, conditions, and cycles used to synthesize the probes are listed in Appendix 9. PCR products were electrophoresed on 1% agarose gels (Gibco), stained in ethidium bromide, and the bands were excised, purified with the GenElute™ Gel Extraction Kit (Sigma) and eluted in 30 to 100 µl

ddH_20. 100 ng of each probe was synthesized with ^{32}P-α-dCTP (10 Ci/ml) using the REDIPrime Kit (Amersham Pharmacia Biotech). 3 filters or $1.5x10^5$ pfu were hybridized with 100 ng of labeled probe in 21 mls of ExpressHyb solution (ClonTech) at 60°C for 2 hrs, followed by washing twice with 2XSSC,0.1% SDS at room temperature, and twice with 0.1XSSC,0.1% SDS at 50°C. Filters were wrapped in saran wrap and exposed with X-OMAT x-ray films for an appropriate time. Up to 35 positive plaque forming units were picked from each plate using a sterile glass Pasteur pipette. Five to seven plaque forming units were pooled and stored in 1X dilution buffer (10 mM $MgSO_4.7H_2O$, 25 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.01% gelatin). Phage were allowed to dissociate overnight at 4°C before converting to plasmids: one third of the diluted phage solution was transfected into *E.coli* strain BM25.8 (grown in LB/10mM $MgSO_4$, OD=0.2). The transfection was done by incubating at 31°C for 30 minutes, then the mixture was incubated in LB broth for another hour shaking (225 rotations per minute) at 31°C. Approximately 10 µl of the transfected *E.coli* strain BM25.8 cells were grown in LB/carbenicillin (50 µg/ml), plasmids were isolated with the GenElute™ Plasmid Purification Kit (Sigma), and eluted in 100 µl ddH_2O. The plasmids were tested for positives using standard PCR with primer pairs and PCR conditions and cycles listed in Appendix 9. Remaining transfected *E. coli* strain BM25.8 cells were plated on 150 mm LB/carbenicillin (50 µg/ml) plates and grown overnight at 37°C for subsequent colony lifting. Colony lifts were performed on lawns of bacterial colonies as for plaque lifts. Some colony lifts were also performed without chemically treating the cells on the filter, by autoclaving up to 2 minutes to lyse the cells. The colony lifts were hybridized with the probes as described above to perform secondary screening. Up to 6 positive colonies were picked and grown in 3 ml LB/carbenicillin. The Plasmids were isolated with GenElute™ Plasmid Purification Kit (Sigma) and eluted in 100 µl ddH_2O. Tertiary screening was performed by PCR using primer pairs and PCR cycles and conditions listed in Appendix 9. Positive plasmids were sequenced with IRD700/800 M13 labeled primers using the Thermo Sequenase Fluorescent Labeled Primer Cycle Sequencing Kit with 7-deaza-dGTP (Amersham Pharmacia Biotech). The reactions were run on a Licor DNA Sequencer Long reader 4200. The sequences were subjected to BLAST searches against high throughput genomic sequence and non-redundant databases of Genbank to check mapping to chromosome 14q13.

4.2.2. Genomic structure determination

The partial cDNA sequences from 5' and 3' RACE, screening of the λ TriplEx cDNA library, and ESTs from GenBank were assembled into consensuses using the GeneTool Version 1.0 software. The consensus sequences, including all possible elucidated alternative transcripts, were subjected to BLAST searches against the non-redundant and high throughput genomic sequence databases of Genbank (E=0.00001, word size=7, mask for low complexity and human repeats, organism=Homo sapiens). The gaps in the BAC clone sequences encompassing each gene were filled in with human genomic sequence from Celera Genomics to construct a composite ungapped genomic sequence. The exon-intron boundaries ascertained from BLAST2 searches were examined to determine whether they were similar to the consensus exon-intron boundary sequences.

In order to determine whether exons 1 and 2 of *NPAS3* were present in the same transcript, RT-PCR was done on human fetal brain Poly A RNA (20-33 weeks gestation, 10 pooled Caucasian male and female) (ClonTech) using methods described in Section

3.2.7. The following primer pairs (f-ggggagagaggcaaaaagtaa;r-gctggccttaaaaacattcg) were used in standard 20 µl PCR reactions. The PCR cycle was denaturation at 94°C-3 mins, followed by 35 cycles of 94°C-30 sec, 57°C-30 sec, 72°C-60 secs, and a final extension of 72°C-5 mins.

4.2.3. Predicted functions by bioinformatics

cDNA sequences were examined for RNA stability signals and Kozak consensus translation start sequences (Kozak, 1996), and examined for potential open reading frames with the GeneTool Version 1.0 software. The following databases available via ExPASY (ca.expasy.org/tools/#proteome) were used to analyze the predicted protein sequences: - physical-chemical properties (ProtParam), pattern and profile searches (InterPro Scan, Coils), cellular localization (PSORTII), post translation modifications (NetPhos), fold family (3D-PSSM), and protein identity (BLASTp, PSI-BLAST).

4.2.4. Construction of enhanced green fluorescent protein plasmids

20 µl PCR reactions were setup with 1X Platinum Pfx Taq buffer, 0.2 mM dNTP, 2 mM $MgCl_2$, 0.4 unit of Platinum Pfx Taq polymerase (InVitrogen), 50 ng of each primer and 2 µl cDNA library template, in order to clone the open reading frames of *NPAS3*, *C14ORF11* and *SNX6* into pCR4-TOPO TA vectors (InVitrogen). The cDNA libraries were λ TriplEx Human Fetal Brain 5' Stretch Plus cDNA Library (pooled 10 male/female, 20 to 25 weeks gestation) (ClonTech) or Marathon-ready human fetal brain cDNA library (21-30 weeks gestation, 10 pooled Caucasian male and female) (ClonTech). The PCR primers and cycles are listed in Appendix 10. The PCR products were electrophoresed on 1% agarose gel (Gibco) stained in ethidium bromide, excised, purified with the GenElute™ Gel Extraction Kit (Sigma), and eluted in 30 µl ddH_20. The

PCR products were poly A tailed and cloned into pCR4-TOPO TA vectors (InVitrogen) as instructed by the manufacturer. Four single colony isolates were selected to harvest plasmids with the GenElute™ Plasmid Purification Kit (Sigma). Plasmids were sequenced with the IRD700/800 M13 labeled primers using the Thermo Sequenase Fluorescent Labeled Primer Cycle Sequencing Kit with 7-deaza-dGTP (Amersham Pharmacia Biotech). Reactions were run on a Licor DNA sequencer Long reader 4200. Cloned open reading frames of *NPAS3* and *C14ORF11* were subjected to site directed mutagenesis using a PCR based method to introduce restriction sites to permit cloning into the EGFP-N1 vector (ClonTech). For *SNX6*, EcoRI restriction sites of the pCR4-TOPO TA vector were used to remove the insert and clone the *SNX6* open reading frame into the EGFP-N1 vector. Each 20 µl PCR reaction was setup with 1X Platinum Pfx Taq buffer, 0.2 mM dNTP, 2 mM $MgCl_2$, 0.4 unit of Platinum Pfx Taq polymerase (InVitrogen), 50 ng of each primer and 2 µl plasmid template. PCR primers and cycles are listed in Appendix 10. The PCR product of *C14ORF11* was poly A tailed and cloned into the pCR4-TOPO TA vector (InVitrogen); *NPAS3* was cloned into pCR-BluntII TOPO vectors (InVitrogen) as described by the manufacturer. Four single colony isolates were selected to harvest plasmids with the GenElute™ Plasmid Purification Kit (Sigma). The plasmids were sequenced with the IRD700/800 M13 labeled primers using the Thermo Sequenase Fluorescent Labeled Primer Cycle Sequencing Kit with 7-deaza-dGTP (Amersham Pharmacia Biotech). Reactions were run on a Licor DNA sequencer Long reader 4200. To clone into the EGFP-N1 vector, the following restriction enzymes were used: *C14ORF11* (EcoRI, BamHI), *NPAS3* (BglII, BamHI) and *SNX6* (EcoRI). 3 µg of the EGFP-N1 vector (ClonTech) was concurrently digested with the appropriate restriction enzyme as instructed by the manufacturer. The single EcoRI digested EGFP-N1 vector was dephosphorylated with two units of calf intestinal phosphatase (Promega) as instructed by the manufacturer before used for cloning. The restricted plasmids were electrophoresed on a 1% agarose gel (Gibco) stained in ethidium bromide; insert and vector DNA were purified with the GenElute™ Gel Extraction Kit and eluted in 30 µl ddH_2O. Each ligation reaction contained 1X T_4 ligase buffer (New England Biolabs), 500 units of T_4 ligase (New England Biolabs), and 3:1 ratio (insert:vector) DNA. The ligation reaction was performed overnight at 16°C. One tenth of the ligation reaction was transformed into DH5α or TOPO F chemically competent cells (InVitrogen) and plated on LB/Kanamycin (30 µg/ml) plates. Four to twenty clones were picked to harvest plasmids with the GenElute ™ Plasmid Purification Kit (Sigma). Two µg of each plasmid was checked for inserts by standard PCR before being sequenced using the EGFPN-1 primers and PCR cycle listed in Appendix 10. A selection of plasmids with inserts were sequenced with the IRDye 800 Terminator Sequencing Kit using the EGFP-N1 forward and reverse primers, in addition to internal primers within the cDNAs. The sequencing cycle was: 94 °C for 5 minutes denaturation, 94°C - 30 sec, 57°C - 30 sec, and 72°C - 60 sec, for 45 cycles. The reactions were run on a Licor DNA Sequencer Long reader 4200.

4.2.5. RT-PCR analysis on cell lines used for transfections

Total RNA was prepared from 1×10^7 to 1×10^8 COS-1 and transformed human skin fibroblasts cells using the RNeasy Miniprep Kit (Qiagen). 1 to 2 µg of total RNA was digested in 1X DNAse I buffer (Gibco) and 2 units DNAse I (Gibco) for 15 minutes at room temperature, then treated with 1.25 mM EDTA for 10 minutes at 65°C to stop the reaction. First strand cDNA synthesis was performed as mentioned in section 3.2.7. The RT-PCR products were diluted to 140 µl with distilled water. 2 µl was used for PCR analysis. Each PCR reaction included as templates: water (negative control), -RT template, or +RT template. The primer pairs used for PCR analysis were: *C14ORF11* (acaacagcctgttccaaatagtgatgct/ aatggcttgctaaaacattgaaaaaatga), *SNX6* (accgcggacttaaagcaataaatgtaga/ tcctcgcacactcaaatcttgattatattcc), *NPAS3* (ggtcagtgtgtgacaaagtactatcgct/ tcaaagtccgagtgctcgaagctg). Each 20 µl PCR reaction contained 0.2 mM dNTP, 1.5 mM $MgCl_2$, PCR buffer (Sigma), 50 ng of each primer, and 0.5 units Taq polymerase (Sigma). For *C14ORF11* and *SNX6*, the PCR cycle was 94°C for 3 minute denaturation, 35 cycles of 94°C-30 sec, 57°C - 30 sec for 35 cycles, followed by 72°C extension for 5 minutes. The PCR cycle for *NPAS3* was 94°C for 3 minute denaturation, 35 cycles of 94°C-30 sec, 52°C - 30 sec, followed by 72°C extension for 5 minutes. The PCR products were analyzed on 1% agarose gel (Gibco) stained in ethidium bromide.

4.2.6. Transfections

22x22 mm cover slips were placed in 6 well plates and sterilized with UV for 2 hrs before use. About 1 x 10^4 to 1 x 10^5 cells/ml were used for each transfection. COS1 and transformed human skin fibroblast cells were grown on cover slips at 37°C in Dulbeccos modification of eagles medium /10% fetal bovine serum 1 to 2 days before transfection. The medium was replaced with 2 mls of fresh Dulbeccos modification of eagles medium/10% fetal bovine serum. To setup the transfection, 3 µl of Fugene 6 reagent was added to Dulbeccos modification of eagles medium/10% fetal bovine serum, then 1 µg of EGFP plasmid construct was added to a final volume of 100 µl. The mixture was incubated for 15 minutes at room temperature before added dropwise to the wells. Transient transfection was allowed to proceed for 24 to 48 hrs at 37°C/5%CO_2. Coverslips with transfected cells were washed in PBS (room temperature) and air dried. 15 μl of mounting medium containing DAPI (Vector Laboratories Inc.) were placed on Fisherbrand Superfrost/Plus slides. The coverslips were placed on the slides and let to sit for 1 hr before being examined with DAPI (to detect the nucleus) and FITC (to detect the EGFP signal) filters of an Olympus BX50 compound microscope (with URA fluorescence). Images were captured with ImageGear 6.6.4 and the software Spot version 2.2.

4.2.7. In silico analysis of gene expression

The assembled cDNA sequences of two genes (*NPAS3*, *C14ORF11*) and *SNX6* from Genbank, were queried against the EST databases of Genbank using BLAST searches (organism=Homo sapiens, E=0.0001, mask for repeats and low complexity regions). Only those alignments between the query sequence and EST hits with high percentage identity (≥ 97%) and full alignment were considered identical. The cDNA library from which the transcript was isolated, and the abundance of the transcript were recorded. Those transcripts isolated from pooled cDNA libraries were ignored.

4.2.8. Human northern analysis

The cDNA isolated from the λ TriplEx 5' Stretch Plus cDNA library for *NPAS3* and *C14ORF11* were used in northern analysis. A 5' RACE product isolated from the Marathon-ready human fetal brain cDNA library and specific to the smaller alternative transcript of *NPAS3* was also used for northern analysis. A probe for *SNX6* was constructed using standard PCR with each PCR reaction containing 2 μl Marathon-ready human fetal brain cDNA library (ClonTech) and 50 ng of the *SNX6* primers (f-cgaacaaagtccaccaatcc; r-ccaaccacgcaaccttaact). The PCR cycle was denaturation at 94°C- 3 mins, followed by 35 cycles of 94°C-30 sec, 58°C-30 secs and 72°C-30 secs, and a final extension of 72°C-5 mins. The probe was electrophoresed on a 1% agarose gel stained in ethidium bromide and gel purified with the GenElute Gel Extraction Kit (Sigma) before being sequenced or radiolabeled. The probe was sequenced with the IRDye 800 Terminator Sequencing Kit using the forward and reverse *SNX6* primers. The sequencing cycle was: 94 °C for 5 minutes denaturation, 94°C - 30 sec, 57°C - 30 sec, and 72°C - 60 sec, for 45 cycles. The reactions were run on a Licor DNA Sequencer Long reader 4200. The human multi tissue northern blots - 12 multi tissue, Blot IV and Blot II (ClonTech) - were prehybridized in 10 ml ExpressHyb solution (ClonTech) at 68°C for 30 mins to 1 hr before adding 10 ng/ml of probe. Probes were labeled with 5 µl ^{32}P-α-dCTP (10 mCi/ml) using the REDIprime Labeling Kit (Amersham Pharmacia Biotech). Hybridization was done for 1 to 2 hrs at 68°C and washes were done twice in 2XSSC, 0.05% SDS for 20 mins at room temperature and twice in 0.1XSSC, 0.1%SDS for 20 mins at 50°C. The blots were wrapped in saran wrap and exposed with Biomax MS or X-OMAT AR X-ray films for an appropriate time.

4.2.9. Extraction of total RNA from mouse embryos

Pregnant mice of the C57BL/6J/AH/BM strain were sacrificed at 6, 7, 8, 9, 10, 11, 12, 14, 16 and 18 days of gestation by cervical dislocation. The embryos enclosed in the deciduas were excised and placed in Dulbeccos modification of eagles medium/4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid for the removal of the remaining deciduas and extraembryonic tissues with a pair of extra-fine pointed splinter forceps. The embryos were then transferred with blunt-pointed forceps to cryovials containing 1 ml of RNAlater solution (Ambion) and stored overnight at 4°C. The heads of 14, 16 and 18 day old embryos and the whole embryos aged at 6 to 12 days were harvested. A maximum of four heads or 10 whole embryos were stored in 1 ml of RNAlater solution (Ambion). Embryos were removed from the RNAlater solution (Ambion) before flash freezing in liquid nitrogen for storage at -80°C. About 200 to 500 mg of pooled whole embryos or embryo heads at a specific developmental stage were crushed into powder with a mortar and pestle prior to using the RNAeasy Miniprep Kit (Qiagen) for total RNA extraction. 2 µg of total RNA were analyzed on 1% formaldehyde agarose gels (1 g agarose, 20 mM 3-[N-morpholino]propanesulfonic acid], 5 mM sodium acetate, 1 mM EDTA, pH 7.0) stained in RNase-free ethidium bromide, to check the integrity of RNA.

4.2.10. RT-PCR analysis of mouse embryonic total RNA

2 µg of total RNA from 6, 7, 8, 9, 10, 11, and 12 day old whole embryos, and heads of 14, 16 and 18 day old embryos were digested with DNAse I and used to synthesize first strand cDNA as mentioned in section 3.2.7. 2 µl of the RT products were used for subsequent PCR analysis with the following primer pairs: *C14orf11* (tgttcaactgaagtggcagtct/ cacagagctgggcagttacc), *Npas3* (aacgcgcatgctgttaactt/ gagaagagtccctcggctgt), *Gapdh* (tgacatcaagaaggtgaag/tccttggaggccatgtaggcc), *Snx6* (tgcatcctgttcccttacttc/ aaaacgccactgctgtctct). These primer pairs were designed with Primer Premiere 3 using mouse orthologous sequences obtained by BLAST searches of these three candidate human cDNA sequences against the non-redundant and EST databases of Genbank (organism= Mus musculus, E=0.0001, mask for repeats and low complexity regions). The GenBank accession numbers were obtained for the following

homologous murine cDNA sequences: NM025456.1 (*C14orf11*), BC025911.1 (*Snx6*) and NM013780.1 (*Npas3*). The template for each PCR reaction contained water only (negative control), -RT template or +RT template. Each 20 µl PCR reaction contained 1.5 mM $MgCl_2$, 1X PCR buffer (Sigma), 1 µM primer, 0.2 mM dNTPs and 0.2 units of AmpliTaq (Sigma). The PCR amplification cycles were: denaturation at 94°C for 5 minutes, 30 cycles of 94°C - 30 sec, 58°C - 30 sec and 72°C - 10 sec, followed by 72°C extension for 5 minutes. The PCR products were analyzed on 1% agarose (Gibco) gels stained in ethidium bromide.

4.2.11. Construction of probes for in situ hybridization

The human cDNA sequences for *C14ORF11*, *NPAS3* and SNX6 were subjected to BLAST searches against the mouse EST database of GenBank to isolate homologous murine cDNA sequences. The GenBank accession numbers were obtained for the following homologous murine cDNA sequences: NM025456.1 (*C14orf11*), BC025911.1 (*Snx6*) and NM013780.1 (*Npas3*). Amplimers were designed from the cDNA sequences using Primer Premiere 3. The amplimers: acaacagcctgttccaaacagtgatgct/ aatggctcgctaaaacattgaaaaaatga (*C14orf11*), accgcggacttaaagcaataaatgtaga/ tcctcgaacactcaaatcttgatgatattcc (*Snx6*), ggtcagtgtgtgaccaagtactaccgct/ tcaaagtccgagtgctcgaagctg (*Npas3*) were used to clone by RT-PCR the mouse homologous sequences for these three genes. Each 20 µl RT-PCR reaction contained 1.5 mM $MgCl_2$, 1x PCR buffer (InVitrogen), 50 ng of each primer, 0.2 mM dNTPs, 0.2 units of Platinum Pfx DNA polymerase (InVitrogen) and 2 µl day 18 mouse RT-PCR DNAse I treated template. The PCR amplification cycles for *C14orf11* and *Snx6* were: denaturation for 3 min at 94°C, 35 cycles of 94°C-30 sec, and 57°C-30 sec, followed by extension at 72°C for 5 min. The PCR cycle for *Npas3* was denaturation at 94°C for 3 mins, 35 cycles of 94°C-30 sec, and 52°C-30 sec, followed by extension at 72°C for 5 min. The PCR products were electrophoresed on 1% agarose gel (Gibco), purified using the GenElute™ Gel Extraction Kit (Sigma) and eluted in 30 μl water. 10 μl of the purified product was poly A tailed and cloned into pCR4-TOPO TA vectors (InVitrogen) as instructed by the manufacturer. Four white colonies were selected to harvest plasmids using the GenElute™ Plasmid Miniprep Kit (Sigma). Plasmids were sequenced with the IRD700/800 M13 labeled primers using the Thermo Sequenase Fluorescent Labeled Primer Cycle Sequencing Kit with 7-deaza-dGTP (Amersham Pharmacia Biotech). The reactions were run on a Licor DNA sequencer Long reader 4200 to confirm clone identity and to ascertain the transcript orientation with respect to T7 and T3 RNA polymerase primer binding sites.

4.2.12. Test of sense and antisense probe expression

Five μg of total RNA from day 9, 10, 11 and 12 gestation embryos (C57BL/6J/AH/BM mice strain) were spotted onto Hybond N^+ filters. The filters were dried at 80°C for 2 hrs before soaking in 2XSSC for 5 mins at room temperature. Plasmids used to synthesize probes were first linearized with 2 μl enzyme, with or without 1X BSA, 1X restriction enzyme buffer and 1 μg template for 2 hrs at 37°C. The restricted products were electrophoresed on 1% agarose gel (Gibco) stained in ethidium bromide, purified with the GenElute™ Gel Extraction Kit (Sigma) and eluted in 30 μl water. Probes were synthesized by asymmetric PCR with each 50 μl reaction containing 5 μl ^{32}P-α-dCTP (10 mCi/ml), 0.08 mM dNTP, 1X PCR buffer (InVitrogen), 200 ng T3 (attaaccctcactaaag) or T7 (taatacgactcactataggg) primer, 0.2 units AmpliTaq (Sigma) and 1 μg template. The

PCR cycle was 94°C- 3 min, followed by 60 cycles of 94°C-20 sec, 44°C-20 sec and 72°C-45 sec. The dot blots were prehybridized with 7 mls of ExpressHyb solution (ClonTech) at 68°C for 30 mins to 1 hr prior to adding the probes and hybridizing for another 1 to 2 hrs at 68°C. Blots were washed in 2XSSC, 0.05% SDS for 20 mins twice at room temperature, then twice in 0.1XSSC, 0.1%SDS for 20 mins at 50°C. The blots were wrapped in saran wrap and exposed in phosphorimager cassettes for up to 3 days. The phosporimager screens were scanned with a Typhoon 8600 variable mode imager (Molecular Dynamics). For *Npas3* and *Snx6* probes, plasmids were linearized with the Pme I (New England Biolabs) restriction enzyme and the T7 primer was used to create sense probes. The plasmids were linearized with the Not I (Gibco) restriction enzyme and the T3 primer was used to create antisense probes. For the *C14orf11*, the plasmid was linearized with the Pme I (NEB) restriction enzyme and the T3 primer was used to create sense probes. The plasmid was linearized with the Not I (Gibco) restriction enzyme and the T7 primer was used to create antisense probes.

4.2.13. Production and spot testing of Riboprobes

1 μg of plasmid was linearized with the PmeI (New England Biolabs) or NotI (Gibco) restriction enzymes and purified as described in section 4.2.12. For *Npas3* and *Snx6* probes, the plasmids were linearized with the Pme I (New England Biolabs) restriction enzyme and the T7 RNA polymerase was used to create sense probes. The plasmids were linearized with the Not I (Gibco) restriction enzyme and the T3 RNA polymerase was used to create antisense probes. For the *C14orf11* probe, the plasmid was linearized with the Pme I (New England Biolab) restriction enzyme and the T3 RNA polymerase was used to create sense probes. The plasmid was linearized with the Not I (Gibco)

restriction enzyme and the T7 RNA polymerase was used to create antisense probes. Riboprobes were made in 50 μl reactions containing 1X transcription buffer (Promega), 7.1 mM DTT, 60 units RnaseOUT (InVitrogen), NTP mix (0.8 mM CTP, 0.8 mM GTP, 0.8 mM ATP, 0.6 mM UTP, 1.25 ng DIG-UTP), 50 units of T3 or T7 RNA polymerase, and two μg linearized plasmid. The reactions were incubated at 37°C for 2 hrs and the DNA was degraded by adding 2 units of DNAse I (Gibco) and incubating the reaction for another 10 minutes at 37°C. The reaction was cleaned with DNase-RNase free NucAway™ spin columns (Ambion) as specified by the manufacture. Hybond N^+ filters (Amersham Pharmacia Biotech) were spotted with water only (negative control), 0.01 ng DIG-UTP (positive control), and 5 μl of each DIG-labeled riboprobe. The filters were baked at 80°C for 2 hrs before use for chemiluminescence detection. To detect the DIG-labeled riboprobes by chemiluminescence, the membranes were washed in 1X TBS (0.05M Tris-HCl, 0.15M NaCl, pH 7.5), then incubated in 20 ml of 1X blocking reagent (Boehringer Mannheim) for 30 minutes. The membranes were washed again twice for 10 minutes in 50 ml TBS before adding 5 μl alkaline phosphatase DIG-UTP antibody (Boehringer Mannheim) in 10 ml TBS and 0.1% Tween-20, and incubating for 30 mins at room temperature. A final wash was done twice in 50 ml TBS, then 50 μl of BCIP/NBT stock solution (Boehringer Mannheim) in 10 ml staining solution (0.1 M Tris-HCl, 0.1 M NaCl, 0.05 M MgCl2, pH 9.5) was added. The reaction was stopped with distilled water after the color developed.

4.2.14. Harvesting of mouse embryos for in situ hybridization

Pregnant mice (C57BL/6J/AH/BM strain) at 11 days post coitus (dpc) were sacrificed by cervical dislocation. The embryos were dissected from the deciduas in ice cold PBS.

The placenta, amnion and chorion were removed and the heart and anterior roof plate of the neural tube were punctured with a 26 G 1/2 needle. The embryos were washed briefly in ice cold PBST (PBS with 0.1% Tween 20) before replacing the PBST with 4% paraformaldehyde fixative solution. Fixation was done overnight at 4°C. The embryos were washed at room temperature in PBST twice for 45 minutes then dehydrated in 25%, 50%, 75%, and 100% methanol/PBST series, with each methanol changed twice for 15 minutes. The embryos were rehydrated for 15 minutes in of each 75%, 50% and 25% methanol series of washes and then twice in PBST before used for *in situ* hybridization. All washes were done at room temperature.

4.2.15. Whole mount in situ hybridization on mouse embryos

Embryos aged at 11 dpc were bleached in 4:1 PBST/hydrogen peroxide for up to 1½ hour and washed three times for five minute intervals in PBST. The above steps were done in RNAse-free scintillation vials. The embryos were then transferred to RNase-free 2 ml screw cap microfuge tubes. The embryos were digested with 10 μg/ml proteinase K for 20 to 30 minutes as judged by the yolk sac dissociation, then washed in PBST. Fixation was done in 4% paraformaldehyde/0.1% glutaraldehyde/PBST for 30 minutes at room temperature followed by washing twice with the series- PBST for 5 minutes, 1:1 PBST/hybridization mix (50% formamide, 5XSSC, 5 mM EDTA, 1 mg yeast transfer RNA, 0.1% CHAPS, 2.5 μg heparin) for 15 minutes and hybridization mix for 1 hr at 70°C. DIG labeled riboprobes in hybridization mix were heated for 3 minutes at 90°C then quenched on ice. Fresh hybridization mix was added before addition of denatured DIG-labeled riboprobes at about 1 μg/ml. Hybridizations were done overnight rotating at 70°C. Post hybridization washes were: twice for 30 minutes at 70°C in prewarmed hybridization mix, twice in wash mix (50% formamide, 1XSSC, 0.1% Tween-20) for 30 minutes at 70°C, then 30 minutes at 70°C in 1:1 wash solution:MABT (100 mM maleic acid, 150 mM NaCl, 0.1% Tween-20, pH 7.5). Rinses were then done three times in MABT, then four half hour washes at room temperature were done in MABT. The embryos were then incubated for 1 hr in 80% MABT/20% blocking reagent (Roche), then for another 1hr in 60% MABT/20% heat inactivated sheep serum/20% blocking reagent, and overnight at 4°C in 60%MABT/50 μg levamisole/20%Blocking reagent/20% heat inactivated sheep serum/ 1/5000 diluted anti-DIG antibody (Boehringer Mannheim). Post antibody hybridization washes were: rinsing three times in MABT, washing nine times in MABT/50 μg levamisole for 1 hr intervals, then overnight to two days in MABT/ 50 μg levamisole. Washing was then done twice in NTMT (100mM NaCl, 100mM Tris pH 9.5, 50 mM $MgCl_2$, 0.1% Tween 20)/ 0.5 mg/ml levamisole for 20 minutes. To detect the antibody, the embryos were incubated in NTMT/50 μg of levamisole/1 in 50 diluted BCIP/NBT (Boehringer Mannheim) for 30 minutes to several hrs. Embryos were then rinsed in 1X PBST and washed twice in PBST before fixing overnight in 4% paraformaldehyde/0.1% glutaraldehdye/PBST. Fixed embryos were rinsed once in PBST and washed twice for 10 minutes each in PBST. The hybridized embryos were analyzed with a Leica MZ12 dissecting microscope with an intralux 4000-1 external light source. Images were captured with a DVC 1300 digital video camera and the Northern elite version 6.0 software.

4.3. RESULTS

4.3.1. Isolation of cDNA sequences

Five potential candidate holoprosencephaly genes were identified in Chapter Three, namely *C14ORF19*, *C14ORF10*, *NPAS3*, *C14ORF11* and *SNX6* after performing 5' and 3' RACE (Figure 4-1). The expression profile in human fetal brain was identified in addition to 5' or 3' cDNA sequence extending beyond the primer binding site. Since the cDNA sequence encoding the protein was known for one of these genes (N99916=*SNX6*, GI: 13027619), the cDNA sequence for the remaining four potential genes was isolated in order to ascertain the potential regulatory regions, potential open reading frames, genomic structure and identity. Screening of a total of 1.9 million plaque forming units from a λ TriplEx Human Fetal Brain 5' Stretch Plus cDNA library resulted in the identification of partial cDNAs for three genes, namely *C14ORF10*, *NPAS3* and *C14ORF11*. Efforts failed to obtain a cDNA sequence for *C14ORF19* since the probe was highly related to numerous members of the immunoglobulin binding protein gene family. Nonetheless, a 600 bp partial cDNA for *C14ORF19*, was obtained from 3' RACE (See Appendix 11). An 822 bp cDNA sequence was predicted with the Genscan algorithm for this gene. To compile the cDNA sequences for *C14ORF10*, *NPAS3* and *C14ORF11*, partial cDNA sequence information obtained from 5' or 3' RACE, screening of the λ TriplEx Human Fetal Brain 5' Stretch Plus cDNA library and ESTs from the database were aligned into composite sequences. The resulting alignments produced a 1.3 kb cDNA of *C14ORF11*, two alternative transcripts of the sizes 1.3 kb and 1.8 kb of *C14ORF10*, and two alternative transcripts of the sizes 2.5 kb and 3.4 kb of *NPAS3* (Appendix 11). Analysis of the cDNA sequences of *C14ORF19*, *C14ORF10*, *NPAS3*,

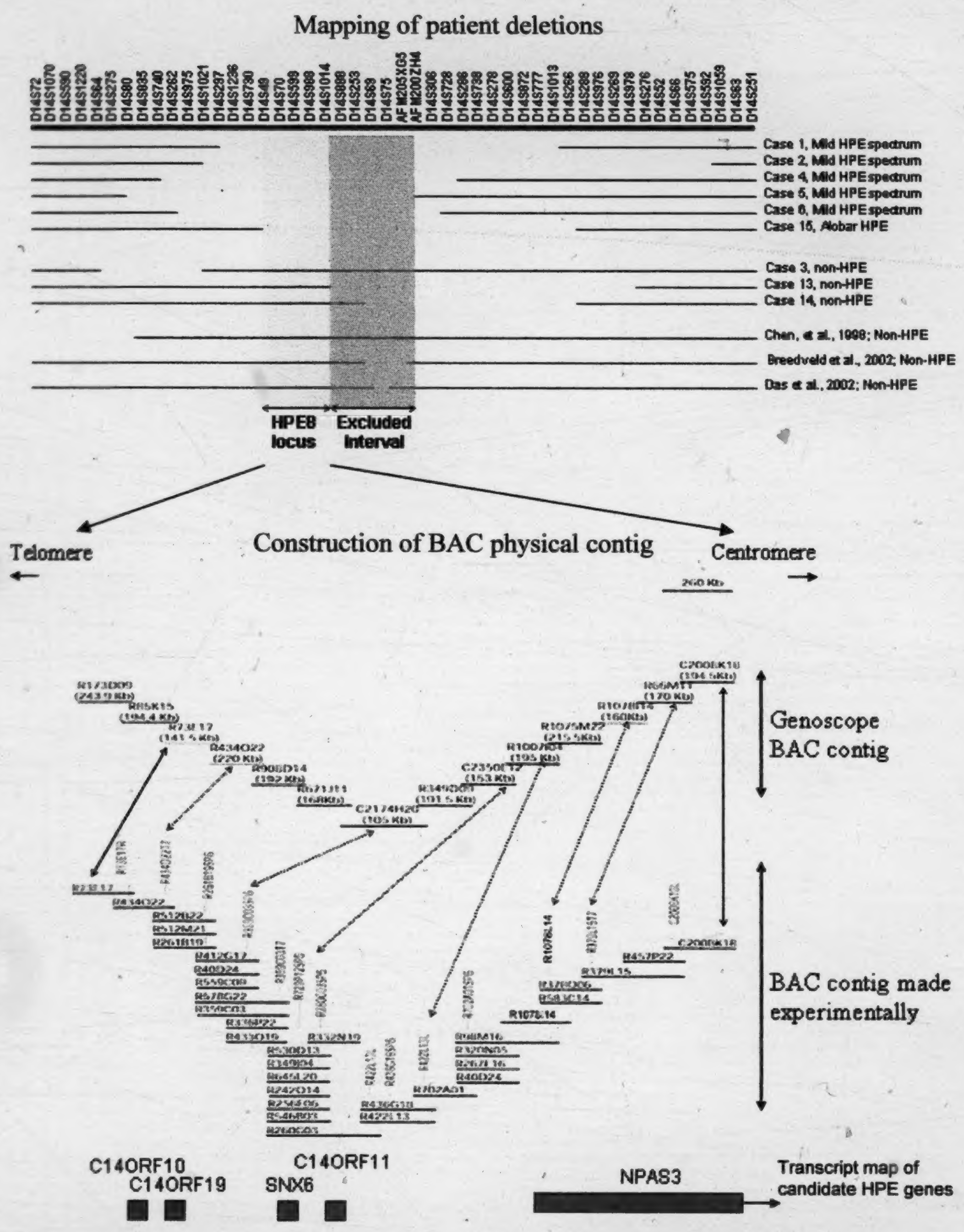
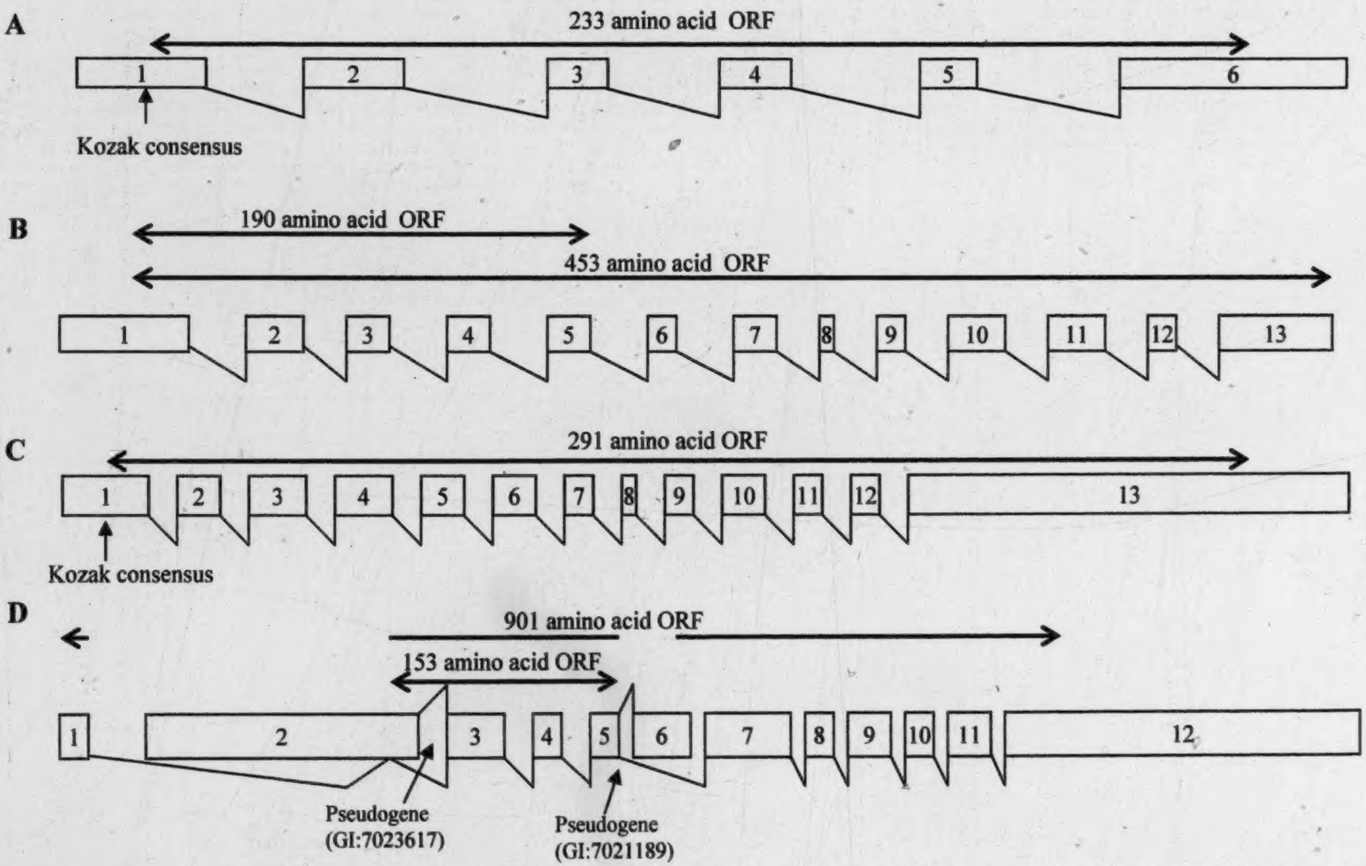

Mapping of patient deletions
D14S72
D14S1070
D14S990
D14S1220
D14S64
D14S275
D14S80
D14S985
D14S740
D14S262
D14S975
D14S1021
D14S287
D14S1236
D14S730
D14S49
D14S70
D14S599
D14S988
D14S1014
D14S888
D14S253
D14S69
D14S75
AFM205XG5
AFM200ZH4
D14S306
D14S728
D14S266
D14S738
D14S278
D14S800
D14S872
D14S777
D14S1013
D14S256
D14S288
D14S976
D14S269
D14S978
D14S276
D14S52
D14S66
D14S575
D14S592
D14S1059
D14S63
D14S251
Case 1, Mild HPE spectrum
Case 2, Mild HPE spectrum
Case 4, Mild HPE spectrum
Case 5, Mild HPE spectrum
Case 6, Mild HPE spectrum
Case 15, Alobar HPE
Case 3, non-HPE
Case 13, non-HPE
Case 14, non-HPE
Chen, et al., 1998; Non-HPE
Breedveld et al., 2002; Non-HPE
Das et al., 2002; Non-HPE
HPE8 locus
Excluded interval
Telomere
Construction of BAC physical contig
Centromere
260 Kb
Genoscope BAC contig
BAC contig made experimentally
C14ORF10
C14ORF19
SNX6
C14ORF11
NPAS3
Transcript map of candidate HPE genes

Figure 4-1: Summary of information on the HPE8 locus. The map positions of five HPE candidate genes are shown below the experimentally built BAC physical contig (anchored to the Genescope contig). Of the five genes found expressed in human fetal brain (20-33 weeks gestation), three (in red boxes) were selected for additional analyses.

C14ORF11 and *SNX6* for regulatory regions showed Kozak consensus sequence was present in the 5' untranslated regions of *C14ORF11* and *SNX6*. None was found for *C14ORF10* and *NPAS3*. Since the 5' end of the *C14ORF19* gene was absent, a Kozak consensus was unlikely to be found. Furthermore, analysis of the 3' ends of all five cDNAs showed that all, including those produced by alternative splicing had polyadenylation signals, (AATAAA) at the 3' end, consistent with a complete 3' sequence of these genes. Since only a partial cDNA sequence was obtained for *C14ORF19*, a related immunoglobulin binding protein gene, and unlikely HPE candidate, no further characterization of this gene was performed.

4.3.2. Determination of the genomic structures

Since both draft Golden Path Sequence of chromosome 14 and Celera Genomics chromosome 14 sequence had deletions and other rearrangements, a composite genomic sequence was created in intervals that encompassed the HPE candidate genes by using sequences from each to fill in the gaps. Any other rearrangements such as small duplications or inversions in the genomic sequence, were ascertained from alignment of the cDNAs reported here in addition to others from the non-redundant or EST databases (Genbank). The cDNA sequences of the genes *C14ORF11*, *C14ORF10*, *NPAS3* and *SNX6* were aligned to the genomic sequence using the BLAST2 algorithm to ascertain the genomic structure. The genomic structures of these four genes are shown in Figure 4-2 and in Table 4-1. These genomic structures cannot be accurately compared with the homologous mouse genomic structures due to the current poor draft quality of conserved mouse genomic sequences in this region (Ensembl, December 2002).



A
233 amino acid ORF
1
2
3
4
5
6
Kozak consensus
B
190 amino acid ORF
453 amino acid ORF
1
2
3
4
5
6
7
8
9
10
11
12
13
C
291 amino acid ORF
1
2
3
4
5
6
7
8
9
10
11
12
13
Kozak consensus
D
901 amino acid ORF
153 amino acid ORF
1
2
3
4
5
6
7
8
9
10
11
12
Pseudogene
(GI:7023617)
Pseudogene
(GI:7021189)

Figure 4-2: Genomic structure of HPE candidate genes. A) *C14ORF11* B) *C14ORF10* C) *SNX6* D) *NPAS3*. Not drawn to scale.

Table 4-1: Genomic structure characteristics of four candidate HPE genes

Gene	Exon	Exon position	Exon size	Intron size	3' acceptor site	5' donor site
C14ORF11	1	1-262	262	3279		GAGCAG**gt**gggctc
	2	263-445	182	2554	tatgct**ag**CTCTGA	GAACTG**gt**aaactt
	3	446-542	96	3968	ttgaat**ag**GATCTT	GAGCAG**gt**aagtaa
	4	543-661	118	4550	aaaatt**ag**TACAGG	AAGGGG**gt**aagaac
	5	662-775	113	8110	tgttct**ag**TTACCA	CCAAAG**gt**aatgga
	6	776-1353	577		tctctt**ag**GCATGA	
NPAS3	1	1-114	114	274137		AGAACG**gt**aacact
	2	115-1864	1749	151798	(gatttc**ag**TTTACA)	TAAAAG**gt**aagttt
	3	1865-1942	77	192855	agttat**ag**GTGCAC	TTGCAG**gt**accttt
	4	1943-2033	90	116393	tcctgc**ag**TCCCTG	TCACAA**gt**aagtaa
	5	2034-2209	175	4221	ccctct**ag**GTGGAG	AGCCAG**gt**gggaat
	6	2210-2649	439	54171	gtctcc**ag**TGGTTT	
	7	2650-2769	119	39014	ttcctc**ag**TGGAGT	TATAAG**gt**aagccg
	8	2770-2964	194	3926	ttttac**ag**GTGATT	AAATAG**gt**actttg
	9	2965-3072	107	15337	ctttgt**ag**GATTAG	TGGACT**gt**aagtac
	10	3073-3221	148	3414	ttcgcc**ag**TGCTGA	TCTTAG**gt**atattt
	11	3222-3347	125	2153	ttccac**ag**CAATCC	TTACAG**gt**attatt
	12	3348-5347	1999		ccacac**ag**AGGACA	
SNX6	1	1-204	204	20296		CGCGGA**gt**aagttc
	2	205-310	105	1499	ttttct**ag**CTTAAA	ACAAAG**gt**aaatgt
	3	311-420	109	2312	ttacag**ag**TTCATT	TATATC**gt**aagtat
	4	421-543	122	2124	tgttac**ag**ATTCCA	GGAAGC**gt**aagtga
	5	544-668	124	29912	gtttct**ag**TGAATA	CAAGAT**gt**gagtat
	6	669-765	96	4399	tctttc**ag**TTGAGT	GTAAAG**gt**aagatt
	7	766-872	106	6739	tcttat**ag**GATGTA	ACAAAA**gt**aagatt
	8	873-949	76	4628	ctccat**ag**GTGCTG	ATGCAA**gt**aagtgt
	9	950-990	40	5718	ctttcc**ag**GTTTTT	ACAAGA**gt**aagtac
	10	991-1078	87	7742	cttttc**ag**AAAATA	GCTAAG**gt**aatttt
	11	1079-1239	160	82	ttttat**ag**GATCTC	AACAAG**gt**actgtt
	12	1240-1326	86	4510	tttcac**ag**AACTTA	GCAAAG**gt**agtgtt
	13	1327-3091	1764		tttttt**ag**GGTAAT	
C14ORF10	1	1-412	412	5170		CAAACA**gt**gagttt
	2	413-541	128	5985	gttacc**ag**AAAAAA	TATAGG**gt**aagtag
	3	542-647	105	597	attttc**ag**CTGCCT	TTACAG**gt**aagaac
	4	648-760	112	1583	ctctta**ag**AACTTA	GTGCAA**gt**aagaat
	5	761-1336	575	294	tttctt**ag**GCAATT	(GAAAAG**gt**gattca)
	6	1337-1408	71	7921	attgtc**ag**TTTGGC	GAATCT**gt**aagtgt
	7	1409-1542	133	2489	attttc**ag**GATTTA	GAACAG**gt**aaaaga
	8	1543-1599	56	77	tcttaa**ag**GAAAGA	TTGGAG**gt**aaatgt
	9	1600-1676	76	1376	tccatt**ag**CTAAGG	TTTATG**gt**atgttc
	10	1677-1814	137	3845	ttccca**ag**GCCAGT	GAAATG**gt**agtagt
	11	1815-1953	138	3063	ttctgt**ag**GACTAT	TTTAGG**gt**aagtct
	12	1954-2015	61	2174	aaatac**ag**GCCATA	TCAAGG**gt**tatctt
	13	2016-2321	305		tcttgc**ag**ATGAAA	

Brackets demarcate cryptic splice sites

C14ORF11 was about 45.8 kb in size with six exons (Figure 4-2a, Table 4-1). Predictions of the largest possible open reading frame (ORF) indicated a 699 bp ORF encoded by all six exons. A 5' UTR (untranslated region) of 193 bp was found in conjunction with a 461 bp 3' UTR. A single murine orthologous *C14orf11* cDNA is reported (GI: 13384859). *C14ORF10* was about 46.8 kb in size with 13 exons (Figure 4-2b, Table 4-1). This gene encoded two alternative transcripts. A small 1.3 kb transcript was encoded by exons 1 to 5. The smaller transcript encoded a 570 bp ORF. A 354 bp 5' UTR was present in addition to a 417 bp 3' UTR. A larger 1.8 kb transcript was encoded by exons 1 to 4, part of exon 5, and exons 6 to 13. A 5' donor cryptic splice site was present at nucleotide 97 of exon 5, which resulted in splicing part of this exon to the 3' acceptor splice site of exon 6. This larger transcript encoded a 1359 bp ORF. A 354 bp 5' UTR and 130 bp 3' UTR were found. Both small and large transcripts shared the same 5' UTR. Only a single *C14orf10* murine orthologous cDNA (GI: 10946957) is thus far reported. *SNX6* was about 78.8 kb in size with 13 exons (Figure 4-2c, Table 4-1). This transcript encoded an 870 bp ORF. The 5' UTR was 496 bp in size, while the 3' UTR was 1713 bp in size. *NPAS3* was about 863 kb in size with 12 exons (Figure 4-2d, Table 4-1). The first four introns were unusually large, being over 100 kb in size. Two possible pseudogenes of sizes 2.2 kb and 2 kb mapped to introns two and five, respectively, of this gene. These genes had 3' poly-A tails and were flanked by inverted repeats. Neither potential pseudogene (GI:7023617, 7021189) was identical to any gene in the database. Furthermore, neither pseudogene was found to encode proteins of over 50 amino acids. *NPAS3* encodes two alternative transcripts of sizes 2.5 and 3.4 kb. The larger transcript is encoded by exon one, part of exon 2, exons 3 to 5, and exons 7 to 12.

A cryptic 3' acceptor splice site was present at nucleotide 1502 of exon two, resulting in the splicing of part of exon 2 to the 5' donor splice site of exon 3. The smaller transcript was encoded by exons 2 to 6. RT-PCR analyses on human fetal brain (20-33 weeks, 10 pool Caucasian male/female) failed to find a product that contained exons one and two, suggesting an internal promoter in intron one. The larger transcript encoded a 2703 bp ORF. The 5' UTR was 64 bp in size while the 3' UTR was 625 bp in size. The smaller transcript encoded a 459 bp ORF with a 1670 bp 5' UTR and 401 bp 3' UTR. Only a single murine *Npas3* orthologous cDNA (GI: 7305318) is reported thus far. For all of the above four genes, the intron-exon boundaries complied with intron-exon consensus sequences (Table 4-1).

4.3.3. Predicted properties of the proteins

The predicted encoded protein sequences of the HPE candidate genes are listed in Appendix 12. A 233 amino acid protein was predicted from *C14ORF11* (Table 4-2). This protein, about 26 kDa in size, was not statistically similar to any other proteins in the database, except for its murine ortholog (58% identity, GI: 13384860). A 290 amino acid protein was predicted from *SNX6* (Table 4-2). This protein, about 290 amino acids and 70% identical to its murine ortholog (GI:23111049), was previously characterized as part of a family of nexin sorting proteins (Parks et al., 2001; Ishibashi et al., 2001). A second human SNX6 protein isoform was recently reported (GI: 23111051), with an additional 116 amino acids in the N terminus. This protein which is 70% identical to the smaller SNX6 protein isoform, may be a predicted artifact from a single 5' sequence of a newly isolated cDNA (GI: 23111050). This cDNA (GI: 23111050) has a 156 bp deletion within the first *SNX6* exon (GI: 13027619) resulting in a frameshift. *C14ORF10* encoded two

Table 4-2: Predicted functional properties of four HPE candidate proteins based on bioinformatic analysis

Gene	Name	Protein Identity	Protein isoform	Number of amino acids	MW (kDa)	pI	Cellular localization[1]	Pattern/Profile	Predicted number of phosphorylation sites			Folding pattern
									Ser	Thr	Tyr	
C14ORF10	Chromosome 14 open reading frame 10	Calcium/ Calmodulin kinase family	small	190	22	9.0	Cytoplasmic (56.5)	Coiled coils (a.a 20-40)	3	1	3	none
			large	453	53	5.0	Cytoplasmic (65.5)	Calcium binding EF hand (a.a 282-334)	11	3	6	EF hand of calmodulin
C14ORF11	Chromosome 14 open reading frame 11		N/A	233	26	4.85	Nuclear (56.5)	Coiled coils (a.a 50-100)	11	4	5	none
SNX6	Nexin sorting protein 6	Nexin sorting protein family	N/A	290	33.5	8.75	Cytoplasmic (52.2)	Phox-like domain (a.a 19-54); Coiled coils (a.a 200-240; 255-290)	8	3	2	none
NPAS3	Neuronal PAS3	bHLH-PAS Domain family	small	153	15	4.7	Nuclear (56.5)	PAS domain (a.a 45-113)	13	2	2	aryl hydrocarbon receptor nuclear translocator
			large	901	97	6.1	Nuclear (82.6)	bHLH dimerization domain (a.a 31-72) Bipartite nuclear localization signal (a.a 568-585) PAS domain (a.a 117-183; 289-355) PAC motif (a.a 361-404)	72	13	8	aryl hydrocarbon receptor nuclear translocator

[1] numbers in parentheses refer to % probability of predicted cellular localization

proteins of 190 and 453 amino acids (Table 4-2). The smaller 190 amino acid protein, about 22 kDa, in size is a truncated isoform of the 453 amino acid protein (about 53 kDa in size) except that the last 23 amino acids of the smaller protein are different. The carboxyl terminal 23 amino acids of the smaller protein are encoded by exon 5. Both small and large protein isoforms are similar to the calcium/calmodulin kinase superfamily. Only a single murine C14Orf10 protein (GI: 10946958) is reported thus far. *NPAS3* encoded two proteins of 153 and 901 amino acids (Table 4-2). The smaller 153 amino acid protein (about 15 kDa in size) was a truncated isoform of the larger 901 amino acid protein (about 97 kDa in size). The larger protein was similar to the bHLH-PAS (**b**asic **H**elix **L**oop **H**elix, **P**eriod, **A**ryl hydrocarbon, receptor, **S**ingle Minded) family of transcription factors and was 90% identical to its murine ortholog (GI: 7305319). Only a single murine Npas3 protein (GI: 7305319) is reported thus far.

Based on these bioinformatics findings, only *SNX6*, *NPAS3* and *C14ORF11* were further characterized since *C14ORF10* seemed a more likely candidate for idiopathic basal ganglia calcification for reasons mentioned in the discussion.

4.3.4. Protein localization studies

Preliminary protein predictions for the cellular localizations of C14ORF11, SNX6 and NPAS3 were listed in Table 4-2. Briefly, the SNX6 protein was predicted to be localized in the cytoplasm and the remaining two proteins, were predicted to be localized in the nucleus.

Before cellular localization studies were done with a selection of these proteins tagged in the C terminus with EGFP, RT-PCR analysis was performed to determine whether the endogenous genes of the above three proteins were expressed in COS1 and

6

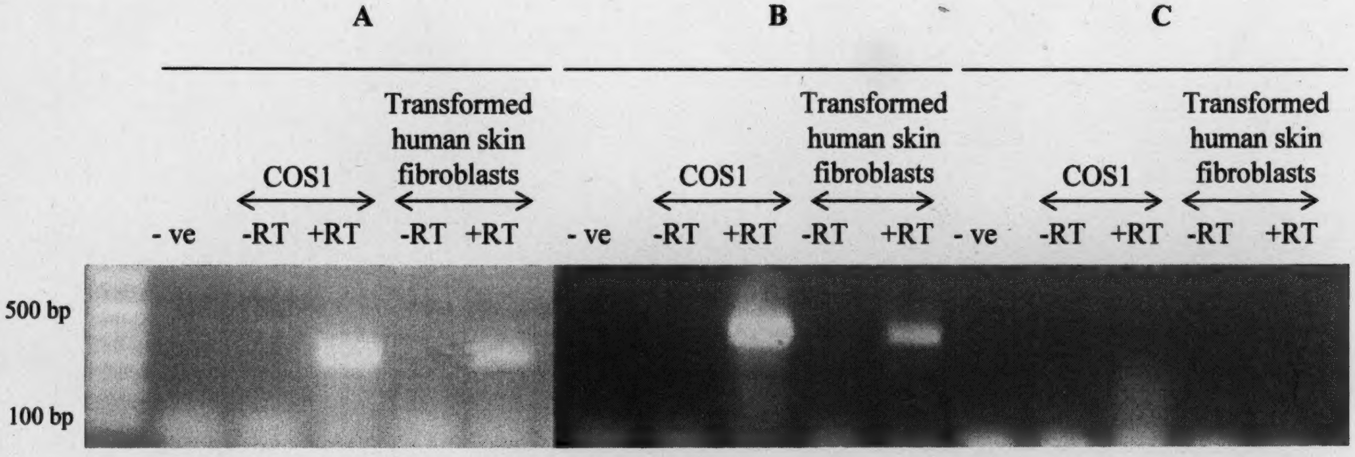

A
B
C
COS1
Transformed human skin fibroblasts
- ve
-RT
+RT
500 bp
100 bp

Figure 4-3: Expression studies of human HPE candidate genes in cell lines by RT-PCR analysis. A) *C14ORF11* B) *SNX6* C) *NPAS3* –ve = water control; -RT = template without reverse transcriptase treatment; +RT = template with reverse transcriptase treament

transformed human skin fibroblast cell lines. Both COS1 and transformed human skin fibroblast cell lines were used to determine the localizations of the EGFP tagged holoprosencephaly candidate proteins. Both endogenous *C14ORF11* and *SNX6* were expressed in COS1 and transformed skin fibroblast cell lines (Figure 4-3 a,b). The endogenous *NPAS3* was not expressed in these two cell lines (Figure 4-3 c).

Only the larger NPAS3 protein isoform was selected for cellular localization studies, on the assumption that the smaller protein may be localized the same as the larger protein (as ascertained from the cellular localization predictions). In addition, due to the small size of the smaller protein isoform, concerns about misfolding when tagged with the EGFP protein were considered. The proteins fused in frame in the C terminus with the EGFP tag when transiently transfected into COS1 or transformed human skin fibroblast cell lines demonstrated that SNX6 and C14ORF11 proteins were localized in the cytoplasm (Figure 4-4 a,b,c,d). NPAS3 protein was localized in the nucleus of both COS1 and transformed human skin fibroblast cell lines (Figure 4-4 e,f). The C14ORF12 protein was not localized in the nucleus as predicted.

4.3.5. Human expression profile studies

Three candidate HPE genes were known to be expressed in human fetal brain (21 to 30 weeks gestation), based on the fact that partial cDNA sequences were isolated from human fetal brain cDNA libraries. Further characterization of the genomic structures showed that for the genes that had alternative transcripts, namely, *NPAS3*, both alternative transcripts were expressed in fetal brain. To determine the expression profile of these three candidate genes in humans, both *in silico* expression and northern analyses were performed. The *in silico* expression pattern of the *C14ORF11* indicated expression

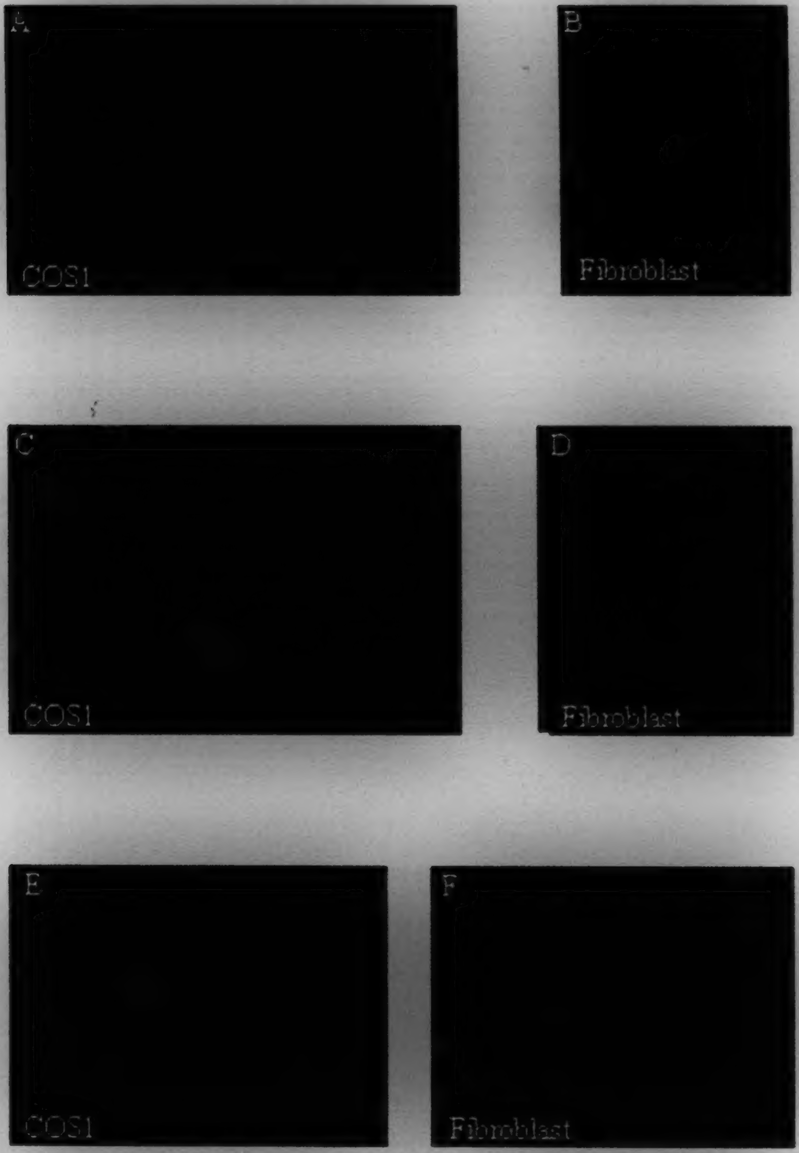

A
COS1
B
Fibroblast
C
COS1
D
Fibroblast
E
COS1
F
Fibroblast

Figure 4-4: Cellular localization studies of HPE candidate proteins. Transient transfections with EGFP-N1 fusion constructs made with (A,B) *SNX6* (C,D) *C14ORF11* (E,F) *NPAS3*. Only EGFP-N1 constructs were made with the 901 amino acid open reading frame of *NPAS3*. Images obtained with DAPI and FITC filters were captured with the Spot version 2.2 software. Blue=DAPI stained nucleus; Green=GFP tagged protein

in fetal and adult tissues in addition to various cancer cell lines (Table 4-3). Northern analysis with cDNA showed a single 1.7 kb transcript that was ubiquitously expressed in several adult tissues including brain (Figure 4-5a). The *in silico* expression pattern of *NPAS3* showed the expression of the 2.5 kb cDNA in placenta and ovarian tumor, while the larger 3.4 kb cDNA expressed in fetal and adult tissues, and several cancer cell lines (Table 4-3). Northern analysis showed an approximately 7.5 kb transcript that was expressed only in various parts of the adult brain using a 2.6 kb probe containing exons 7 to 12 (Figure 4-5b). No expression of the smaller transcript was noted in adult tissues using a 1.6 kb cDNA probe of exon two. The *in silico* expression pattern of *SNX6* showed expression in fetal and adult tissues in addition to several cancer cell lines. Northern analysis with a probe encompassing exons four to eight of the cDNA (GI:13027619) showed a 3.4 kb transcript that was ubiquitously expressed in several tissues including brain (Figure 4-5c).

4.3.6. Mouse embryonic expression profile studies by RT-PCR

The mouse embryonic expression profile from days 6 to 18 of the C57BL/6J/AH/BM strain was used to obtain the temporal range of embryonic expression for the murine homologues of the three HPE candidate genes. RT-PCR primers for the three genes were designed, based on murine sequences in the databases (Genbank) that were highly homologous to the human cDNA sequences of these genes. By using RT-PCR analyses, all three genes were found expressed from days 6 to 12, using total RNA extracted from the whole embryos (Figure 4-6). All three genes were also expressed from days 14 to 18 from total RNA extracted from the heads of the embryos (Figure 4-6).

Table 4-3: *In silico* expression profile of three candidate HPE genes

			In silico expression pattern	
Gene		**Fetal**	**Adult**	**Cancer**
NPAS3	2.5 kb cDNA	Placenta (1)		Ovarian tumor (2)
	3.4 kb cDNA	Placenta (1)	Hypothalamus (2), uterus (1)	Anaplastic oligodendroglioma (3), glioblastoma (3), medulloblastoma (2), ovary tumor (2)
C14ORF11		Hypothalamus (1), eye (2), ear (1), total fetus (3), cochlea (1), placenta (2)	Leukocyte (1), medulla (1), uterus (2), T-lymphocyte (1), erythroid cells (1), prostate (2), bone marrow (1), kidney (7), colon (4), testis (2), breast (1), multiple sclerosis lesions (1), dendritic cells (2), germinal center B cells (1)	Neuroblastoma (1), lymphoma (1), adrenal cortex carcinoma (1), parathyroid tumor (4), lung carcinoma (2), oligo dendroglioma (1), glioblastoma (1), mammary adenocarcinoma (3)
SNX6		Eye (2), placenta (3)	Stomach (3), optic nerve (1), kidney (1), retina (1)	Teratocarcinoma (1), cervical carcinoma (1), neuroblastoma (1), retinoblastoma (1), parathyroid tumor (1)

Numbers in parentheses refer to abundance of transcript source in the human EST database (January, 2002).

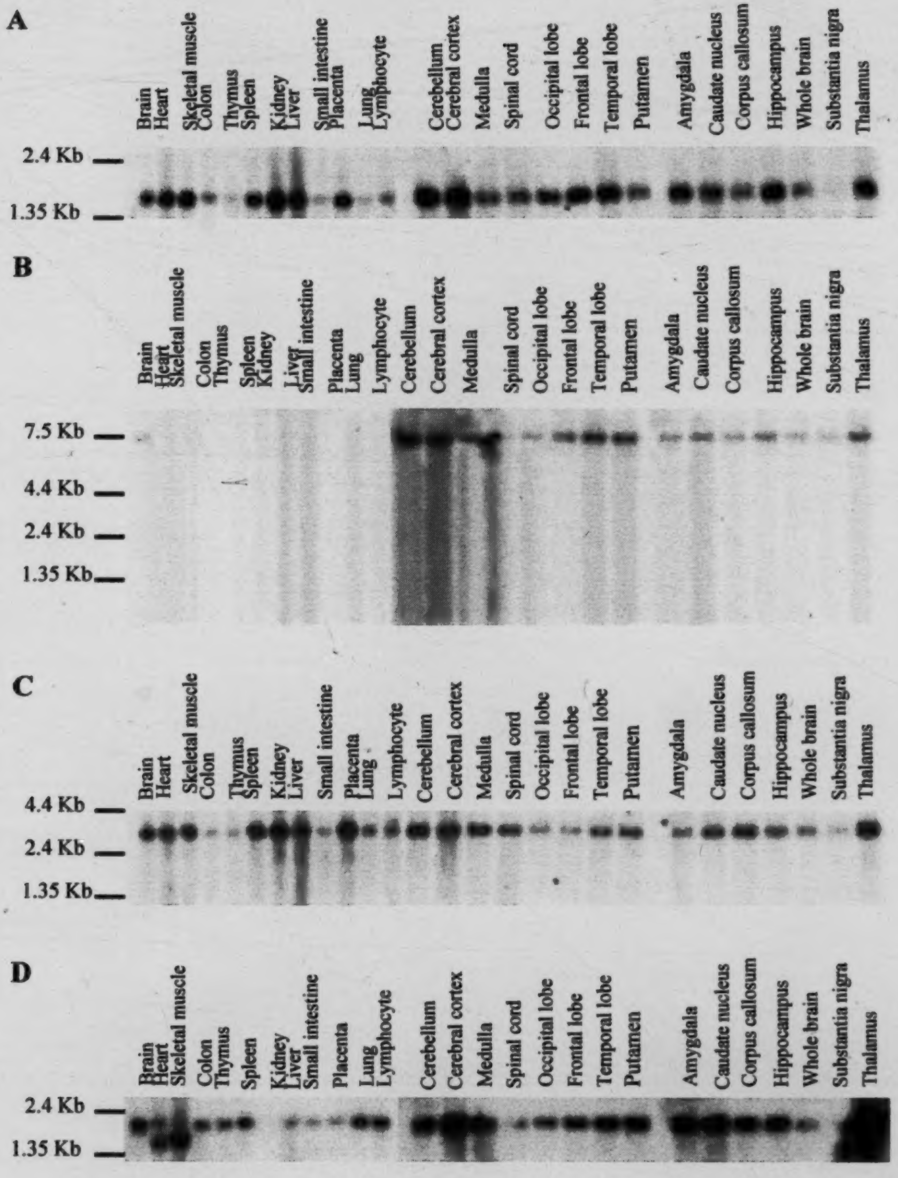
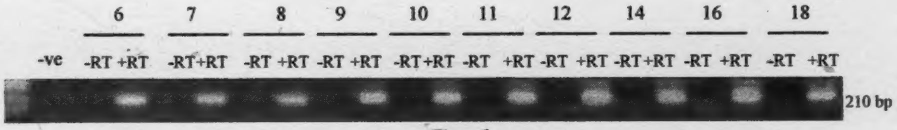
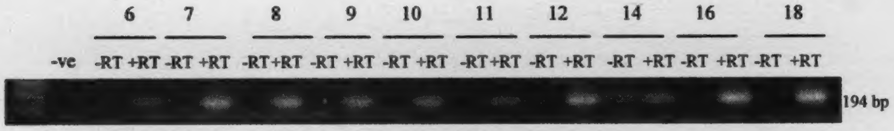

A
Brain
Heart
Skeletal muscle
Colon
Thymus
Spleen
Kidney
Liver
Small intestine
Placenta
Lung
Lymphocyte
Cerebellum
Cerebral cortex
Medulla
Spinal cord
Occipital lobe
Frontal lobe
Temporal lobe
Putamen
Amygdala
Caudate nucleus
Corpus callosum
Hippocampus
Whole brain
Substantia nigra
Thalamus
2.4 Kb
1.35 Kb
B
Brain
Heart
Skeletal muscle
Colon
Thymus
Spleen
Kidney
Liver
Small intestine
Placenta
Lung
Lymphocyte
Cerebellum
Cerebral cortex
Medulla
Spinal cord
Occipital lobe
Frontal lobe
Temporal lobe
Putamen
Amygdala
Caudate nucleus
Corpus callosum
Hippocampus
Whole brain
Substantia nigra
Thalamus
7.5 Kb
4.4 Kb
2.4 Kb
1.35 Kb
C
Brain
Heart
Skeletal muscle
Colon
Thymus
Spleen
Kidney
Liver
Small intestine
Placenta
Lung
Lymphocyte
Cerebellum
Cerebral cortex
Medulla
Spinal cord
Occipital lobe
Frontal lobe
Temporal lobe
Putamen
Amygdala
Caudate nucleus
Corpus callosum
Hippocampus
Whole brain
Substantia nigra
Thalamus
4.4 Kb
2.4 Kb
1.35 Kb
D
Brain
Heart
Skeletal muscle
Colon
Thymus
Spleen
Kidney
Liver
Small intestine
Placenta
Lung
Lymphocyte
Cerebellum
Cerebral cortex
Medulla
Spinal cord
Occipital lobe
Frontal lobe
Temporal lobe
Putamen
Amygdala
Caudate nucleus
Corpus callosum
Hippocampus
Whole brain
Substantia nigra
Thalamus
2.4 Kb
1.35 Kb

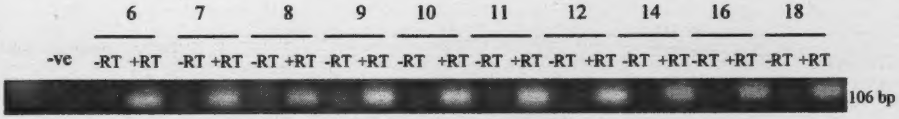
Figure 4-5: Human northern analysis of HPE candidate genes. A) *C14ORF11*; B) *NPAS3*; C) *SNX6*; D) *GAPDH* (Glyceraldehyde-3-phosphate dehydrogenase, control)


6
7
8
9
10
11
12
14
16
18
-ve
-RT +RT -RT+RT -RT +RT -RT +RT -RT+RT -RT +RT -RT +RT -RT+RT -RT +RT -RT +RT
210 bp

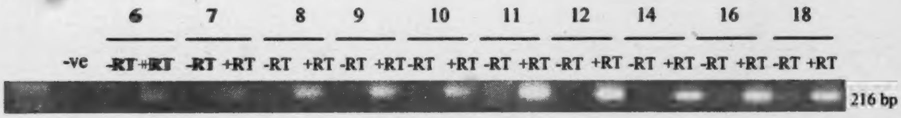
Snx6


6
7
8
9
10
11
12
14
16
18
-ve
-RT +RT -RT +RT -RT+RT -RT +RT -RT +RT -RT+RT -RT +RT -RT +RT -RT +RT -RT +RT
194 bp

Npas3


6
7
8
9
10
11
12
14
16
18
-ve
-RT +RT -RT +RT -RT +RT -RT +RT -RT +RT -RT +RT -RT +RT -RT +RT-RT +RT -RT +RT
106 bp

C14orf11


6
7
8
9
10
11
12
14
16
18
-ve
-RT +RT -RT +RT -RT +RT -RT +RT-RT +RT -RT +RT -RT +RT -RT +RT -RT +RT -RT +RT
216 bp

Gapdh

Figure 4-6: Expression studies by RT-PCR on the HPE8 candidate murine holoprosencephaly orthologues. Murine total RNA was extracted from days 6 to 12 whole embryos; and the head of 14 to 18 embryos. -ve = water control; –RT = template without reverse transcriptase treatment; +RT = template with reverse transcriptase treatment.

4.3.7. Whole mount *in situ* hybridization studies

In order to determine the spatial expression of the murine homologues of these three genes during mouse embryogenesis, whole mount *in situ* hybridization studies were performed. Probes designed for these studies were 1) unique and therefore not part of a conserved sequence motif and, 2) spanned at least two exons such that the heterogeneous RNA would have intervening intron sequences. Whole mount *in situ* hybridization was selected on day 11 since the embryos are easy to harvest at this stage, the protocol was more feasible and neurulation and patterning of the murine brain still occurs during this gestational period. Expression of *C14orf11* was found along the developing spine, central nervous system, face, pharynx, limbs and viscera (Figure 4-7a). Expression of *Snx6* was found along the developing central nervous system, spine, face, limbs and viscera (Figure 4-7b). Expression of *Npas3* was previously reported (Brunskill et al., 1999). The whole mount *in situ* expression profile during 11 dpc was done in order to confirm the previous results. The whole mount *in situ* hybridization results showed expression in the developing central nervous system and spine (Figure 4-7c). For all three genes, these signals were not present with the control (sense) hybridizations (Figure 4-7).

4.4. DISCUSSION

4.4.1. Characteristics of the 14q13 holoprosencephaly candidate genes

Five candidate genes for holoprosencephaly were proposed in Chapter Three based on the fact that they were expressed in human fetal brain (21 to 30 weeks gestation). One of the five candidate genes, *SNX6,* had a cDNA identified and characterized by others (Parks et al., 2000; Teasdale et al., 2001; Ishibashi et al., 2001). The remaining four candidate

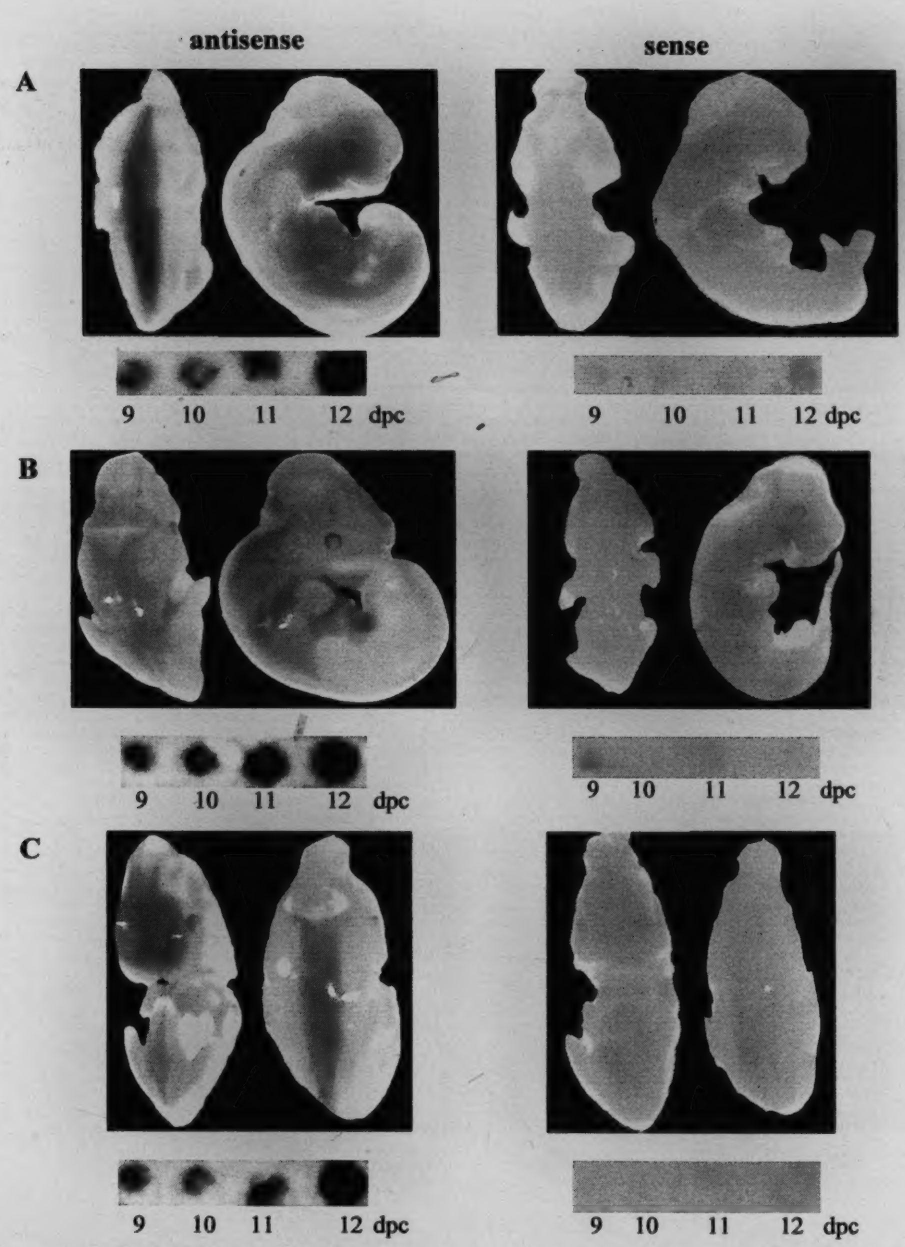

antisense
sense
A
9 10 11 12 dpc
9 10 11 12 dpc
B
9 10 11 12 dpc
9 10 11 12 dpc
C
9 10 11 12 dpc
9 10 11 12 dpc

Figure 4-7: Expression studies by whole mount *in situ* hybridization on selected day 11 C57BL/6J murine embryos. A) *C14orf11*. A 340 bp probe was used spanning exons 5 and 6. B) *Snx6*. A 484 bp probe was used spanning exons 1 to 5. C) *Npas3*. A 448 bp probe was used spanning exons 9 to 12. Map position of probes are discerned from mouse genomic structures (Ensembl, December 2002), and may change depending on the improvement of the genomic structure determination. RNA dot blot analysis with sense and antisense probes are shown below each day 11 embryo. dpc = days post coitus. Of note, the signal variation on the dot blots is due to technical variation in spotting of total poly A RNA.

genes had only a partial EST sequence in the database (Genbank) when this research was in progress. The subsequent assembly of RACE, cDNA library screening and EST sequences resulted in the identification of two alternative cDNAs of *C14ORF10* and *NPAS3* and a single cDNA of *C14ORF11*. For these genes, the cDNA sequence was probably not full length. However, the complete 3' ends of these genes were obtained since they contained polyadenylation signals. A large cDNA could not be obtained for *C14ORF19*, hence this gene was omitted for further analysis. The genomic structures for the remaining four genes were further ascertained. The bioinformatic analysis of these four candidate genes with isolated cDNA sequences led to the exclusion of *C14ORF10* as a possible candidate for holoprosencephaly. Functional characterizations of the remaining genes, namely, *C14ORF11*, *SNX6* and *NPAS3*, were further assessed with respect to putative properties predicted by bioinformatic analysis, cellular localization studies, human *in silico* and adult tissue expression profiles, and embryonic expression studies in the C57BL/6J/AH/BM mouse strain, in order to select the most likely candidates for the cause of the holoprosencephaly phenotype seen among the patients with proximal chromosome 14q deletions.

4.4.1.1. *C14ORF19* (chromosome 14 open reading frame 19)

A 600 bp partial cDNA sequence of *C14ORF19* was obtained. This gene is part of an immunoglobulin binding protein family that functions in B cell signal transduction (Onda et al., 1997). Due to fact that an incomplete cDNA was obtained for this gene, detailed characterization of the gene was not pursued. It is unlikely that a protein of the immunoglobulin binding protein family is a possible candidate for holoprosencephaly.

4.4.1.2. *C14ORF10* (chromosome 14 open reading frame 10)

C14ORF10 is 46.8 kb in size with 13 exons. This gene has two alternative transcripts (1.3 and 1.8 kb) that encode two protein isoforms. Both transcripts are likely produced under the transcriptional control of the same promoter since the 354 bp 5' UTR of each transcript was the same and resided in the first exon. Both proteins were similar to the superfamily of calcium/calmodulin kinases, suggesting as an unlikely candidate for holoprosencephaly. The calcium/calmodulin signaling pathway is implicated in neural physiologic function (Wang and Kelly, 2001), cardiac hypertrophy (Ruwhof and van der Laarse, 2000), skeletal muscle hypertrophy (Semsarian et al., 1999), and Alzhiemer's disease (Muthalif et al., 2002), where it is primarily involved in calcium homeostasis. The role of the calcium/calmodulin signaling pathway in neural development is unknown. This gene is more likely an excellent candidate for idiopathic basal ganglia calcification (Fahr disease) since it maps within 500 kb proximal to the marker D14S1014 that showed linkage in patients with this disease (Geschwind et al., 1999). The age of onset for idiopathic basal ganglia calcification is 30 to 60 yrs. The key hallmarks of this disorder include dysarthria, ataxia, progressive dystonia, parkinsonism and neuropsychiatric manifestations with calcification of the central nervous system.

4.4.1.3. *SNX6* (nexin sorting protein 6)

A 3 kb cDNA sequence of the *SNX6* was isolated by Parks et al., (2001) (GI: 13027619) from a yeast two hybrid screen to identify proteins that interact with SMAD1. When this research was in progress, the genomic structure and functional characterization as outlined in the results were unknown except for the cellular localization in COS1 cells and limited adult human expression pattern. In this research, the genomic structure of the

SNX6 gene was determined to be 78.8 kb in size with 13 exons. This gene was found to be expressed in human fetal brain (21-33 weeks gestation), and a single 3.4 kb transcript was found ubiquitously expressed in several adult tissues including brain, using a probe that spanned exons 4 to 8. Sequencing of the probe confirmed its identity as being part of nexin sorting protein 6 transcript. Previous northern expression analysis by Parks et al. (2001) showed two alternative transcripts of 3 and 2.2 kb sizes that had ubiquitous expression in several adult human tissues, using a 3 kb cDNA probe which also covered exons 4 to 8 of the transcript. The 3.4 kb transcript obtained in this study is more likely another alternative transcript that was undetected in the report by Parks et al. (2001). Furthermore, these findings imply that alternative splicing downstream of exon 8 and upstream of exon 4 produces 3 and 2.2 kb transcripts. A 2.9 kb alternative human SNX6 cDNA (GI: 23111050) has been recently reported, encoding a hypothetical SNX6 protein isoform with 70% identity to that used in this study. This cDNA sequence, which has a 156 bp deletion within the first exon of *SNX6* (GI: 13027619), is more likely an artifact of a deletion produced by the reverse transcriptase. The *in silico* expression profile of this gene also demonstrated expression in several fetal and adult human tissues in addition to several cancer cell lines. The murine orthologue of this gene was shown to express during 6 to 18 gestational days in the C57BL/6J developing mouse embryos. Whole mount *in situ* hybridization on day 11 gestation of C57BL/6J mouse embryos showed expression along the developing spine, central nervous system, face, limbs and viscera.

Cellular localization results showed that the SNX6 protein was localized in the cytoplasm of COS1 and transformed human skin fibroblast cell lines, confirming previous studies which showed localization to the cytoplasm of COS1 and Hela cells (Parks et al., 2001; Teasdale et al., 2001). When associated with the Pim-1 protein, the SNX6 protein can translocate to the nucleus of Hela cells (Ishibashi et al., 2001). In this study, the endogenous *SNX6* gene was expressed in both COS1 and transformed skin fibroblast cell lines. However, it is yet to be determined whether the *PIM1* gene is expressed in the cell lines used for the present cellular localization studies. Since no nuclear localization of the SNX6 protein was detected in this study, lack of *PIM1* gene expression or other missing factors may prevent localization of the SNX6 protein to the nucleus. The C terminal tagging of the SNX6 protein with EGFP may possibly result in structural changes preventing interaction with other proteins required for trafficking to the nucleus or other cellular compartments.

The fact that the SNX6 protein has been found to interact with and traffick the transforming growth factor beta receptor family (TGFß I and II) and transforming growth factor beta receptor serine/threonine kinases (Parks et al., 2001) suggests that it may have dual roles in both trafficking and regulating the nodal receptors that are known to bind to ligands such as cripto, nodal, one-eyed pinhead and Cyclops and transduce signaling via transcription factors such as TGIF (Schier and Shen, 1999; Goumans and Mummery, 200). Some of the latter genes are mutated in HPE patients (TDGF1/CRIPTO, TGIF) (de la Cruz et al., 2001; Gripp et al., 2000; Chen et al., 2002), or result in a holoprosencephaly phenotype in model organisms such as zebrafish (one-eyed pinhead, Cyclops) (Rebaglianti et al., 1999; Feldman and Stemple, 2001). The SNX6 protein has been found to heterodimerize with other SNX protein family members such as SNX1, 2 and 4 (Parks et al., 2001). Other members of the nexin sorting protein family are involved in trafficking other membrane receptors such as epidermal growth factor and platelet-derived growth factor (Parks et al., 2001). The heterodimeric property of SNX6 may also implicate it in the trafficking of these other receptors. Thus far, none of the 15 nexin sorting protein family members is associated with a disease phenotype, except *SNX3* is recently reported to be associated with microcephaly, microphthalmia, ectrodactyly and prognathism (Vervoort et al., 2002). Based on the embryonic expression pattern reported in this study and that fact that the SNX6 protein functions in the transforming growth factor signaling pathway, the *SNX6* gene is an excellent candidate for holoprosencephaly.

4.4.1.4. ***C14ORF11*** (chromosome 14 open reading frame 11)

C14ORF11, 45.8 kb in size with six exons, not similar to any proteins except its murine ortholog with unknown function, is possibly an accessory protein that mediates its function in the predicted coiled-coil domain of the amino terminus. Cellular localization studies showed the protein localized in the cytoplasm of COS1 and transformed human skin fibroblast cell lines, rather than in the nucleus as predicted. The endogenous *C14ORF11* was expressed in both cell types, hence other proteins or co-factors that are required for proper localization may be present. The fact that the protein was not localized to the nucleus could be a result of structural problems from the carboxyl terminal fusions with EGFP, resulting in the inability to interact with trafficking proteins or accumulation in the Golgi. Another possibility is that the protein is not localized in the nucleus in these cell types. *C14ORF11* was expressed in several human adult, fetal and cancer cell lines. The murine *C14orf11* showed expression during 6 to 18 gestational days in C57BL/6J mouse embryos, as well along the developing central nervous system and other organs of day 11 embryos, suggesting that the *C14ORF11* gene is another candidate for holoprosencephaly.

4.4.1.5. *NPAS3* (neuronal PAS3)

This gene about 863 kb with 12 exons encodes two alternative transcripts. The smaller transcript encoded a 153 amino acid protein identical to part of the larger 901 amino acid protein. The larger protein has bHLH-PAS domains that belong to a family of bHLH-PAS transcription factors (Crews and Fan, 1999; Crews, 1998). The smaller protein has a PAS domain that may heterodimerize with the larger protein isoform or even other transcription factors with PAS domains to mediate dominant negative regulation (Crews and Fan, 1999; Crews, 1999). This mode of regulation has been proposed for the PITX2D protein isoform which regulates the three other protein isoforms (Cox et al., 2002). Members of the bHLH-PAS transcription factor family are associated with diverse roles such as neurogenesis, circadian rhythms, toxin metabolism and even tracheal development. One *Drosophila* homolog, namely Single minded (*Sim*), is involved in patterning of the central nervous system midline (Thomas et al., 1988; Nambu et al., 1991). If this gene is mutant, the CNS midline is lost and the ventricles fuse (Thomas et al., 1988; Nambu et al., 1991) in a manner reminiscent of alobar holoprosencephaly. Furthermore, the *Sim* orthologues are also implicated in regulating the expression of *Shh* (Epstein et al., 2000).

Studies for the larger protein showed localization to the nucleus of both COS1 and transformed human skin fibroblast cell lines, as predicted. The endogenous *NPAS3* was, however, not expressed in these cell types implying that the requirements for the trafficking of the larger protein isoform were still present in both cell lines analyzed.

Cellular localization studies were not done with the smaller predicted protein based on the assumption that localization may possibility be the same as the larger protein.

Both transcript isoforms were expressed in human fetal brain (21 to 30 weeks gestation), however only the larger transcript isoform was expressed in adult brain. The murine *Npas3* gene is expressed only in the brain of adult tissues examined (Brunskill et al., 1999). Further expression analyses with the murine *Npas3* demonstrated expression during 6 to 18 gestational days of C57BL/6J mouse embryos. Whole mount *in situ* hybridization on gestational day 11 C57BL/6J mouse embryos showed expression in the developing spine and central nervous system, using a probe specific to the 3' end of the murine *Npas3* gene (GI: 7305318). This probe is homologous to the human *NPAS3* large transcript isoform ascertained in this study, but not the small transcript isoform. Only a single murine *Npas3* cDNA (GI: 7305318), likely the ortholog of the human large *NPAS3* transcript isoform (71% identity), is reported thus far. The embryonic expression profile of the murine *Npas3* was previously studied (Brunskill et al., 1999), using a probe spanning exons 1 to 5. The reported expression pattern during the gestational day 11 was the same as that reported here. No expression was previously reported to be detected before day 9.5 by whole mount *in situ* hybridization (Brunskill et al., 1999). This difference from the present study was likely due to low expression levels which were detected here by RT-PCR. Brunskill et al., (1999) also reported that murine *Npas3* was expressed in the developing neural tube from 9.5 dpc and into the entire neuroepithelium of the central nervous system by 10 to 12 dpc. By 14.5 dpc, expression was restricted to the neopallial layer of the cortex in addition to pericardium, primitive mesenchymal tissue surrounding the nasal cavity, dermal connective tissue below the epidermis and mesenchyme of the limb bud, and developing ear (Brunskill et al., 1999). By day 16.5 dpc, expression was detected in the submandibular salivary gland and epithelial tubules of the kidney (Brunskill et al., 1999). The *NPAS3* gene is another excellent candidate for holoprosencephaly based on its embryonic expression pattern and the fact that other gene family members such as *Sim* (single minded) are implicated in neurogenesis and holoprosencephaly phenotypes.

4.5. SUMMARY OF FINDINGS

1) The cDNA sequences for three genes, namely *C14ORF11*, *NPAS3* and *C14ORF10* were isolated, and a partial cDNA for *C14ORF19* was isolated.

2) The genomic structures for four genes were determined, namely *C14ORF11*, *NPAS3*, *C14ORF10* and *SNX6*. The latter gene had a cDNA reported by another group.

3) Based on the bioinformatics analysis of the five genes initially proposed as holoprosencephaly candidates in Chapter Three, three genes were selected as candidates, namely, *SNX6*, *C14ORF11* and *NPAS3*.

4) Cellular localization studies showed one (NPAS3) of the three HPE candidate proteins was localized in the nucleus, while the two remaining proteins (C14ORF11, SNX6) were localized in the cytoplasm.

5) Human Northern analysis showed ubiquitous expression of *SNX6* and *C14ORF11* in several adult tissues of the body and brain, however *NPAS3* was expressed only in the brain.

6) Mouse embryonic temporal expression profile by RT-PCR analyses showed all three murine orthologues of the HPE candidate genes were expressed during 6 to 18 dpc, furthermore whole mount *in situ* hybridization on 11 dpc murine embryos showed expression in the developing spine, central nervous system and/or other extra-cranial structures.

7) The findings of this chapter have led to the proposal of three excellent candidate genes (*NPAS3, SNX6, C14ORF11*) for holoprosencephaly at 14q13.

CHAPTER FIVE

ANALYSIS OF CANDIDATE GENES FOR REARRANGEMENTS IN SELECTED PATIENTS WITH DERIVATIVE 14 TRANSLOCATION CHROMOSOMES

The following have contributed to work in this chapter:

Metaphase and Fiber FISH analyses on cases 11 and 12 were done in collaboration with Dr. Benjamin Picard with Dr. Walter Muir (University of Edinburgh, UK)

Patricia O'Brien with Professor Ferguson-Smith (Cambridge University, UK) isolated aberrant chromosomes from lymphoblast cell lines of selected patients using flow cytometry

Parts of this chapter have been published in:

1. **Kamnasaran D**, O'Brien PCM, Zackai EH, Muenke M, Ferguson-Smith MA and Cox DW. Rearrangement in the *PITX2* and *MIPOL1* genes in a patient with a t(4;14) chromosome. European Journal of Human Genetics 2003. In press.
2. **Kamnasaran D**, Muir WJ, Ferguson-Smith MA and Cox DW. Disruption of the Neuronal PAS3 gene in a family with schizophrenia. Journal of Medical Genetics 2003. In press.

5.1. INTRODUCTION

Chromosome aberrations involving translocations between the 14q13 cytogenetic band and other chromosomes are relatively rare. Only a few such cases have been reported (Kim et al., 1997; Ohyashiki et al., 1990; Fryns et al., 1984; Wang et al., 2002; Buchanan et al., 1978; the Mitelman database (cgap.nci.nih.gov/Chromosomes/ RecurrentAberrations). In Chapter Two, the definition of translocation breakpoint junctions in six patients and the deletion intervals in nine patients, was described. In Chapters Three and Four, the identification and characterization of holoprosencephaly candidate genes at HPE8 was reported. Of the six translocation cases characterized in Chapter Two, three cases are of particular interest. In Section A, a t(4;14) patient with agenesis of the corpus callosum and other anomalies was examined in detail since agenesis of the corpus callosum was evident in six of the deletion cases reported in Chapter Two. In section B, a mother and daughter who are carriers of a t(9;14) chromosome and affected with schizophrenia , were examined in detail since Chapter Two reported the translocation breakpoint junctions mapped to the proximal interval of the defined HPE8 locus. In this chapter, these three translocation cases were examined for the causative genes on the derivative chromosome 14 and/or the other derivative chromosome in order to explain the phenotypes observed. The premise of this chapter therefore is to test the sub-hypothesis that these "*translocation chromosomes are important in excluding or including genes and disease phenotypes that are associated with a specific chromosome region*". The data of this chapter have identified candidate genes for agenesis of the corpus callosum (Section A) and schizophrenia (Section B) in proximal chromosome 14q translocation cases.

5.2. SECTION A – ANALYSIS OF t(4;14) CASE 10 (Cell line L6404)

5.3.1. The *MIPOL1* gene

The *MIPOL1* (mirror image polydactyly 1) gene, mapping at 14q13, contains 15 exons that span about 350 kb (Kondoh et al., 2002). This gene encodes a 442 amino acid protein that has two coiled-coil domains in the carboxyl terminus, one of which is highly similar to the caspase recruitment domain. The first four exons of the gene are non-coding. The gene has three alternative transcripts of 7, 3 and 2 kb sizes and is expressed in adult kidney, skeletal muscle, liver, heart, pancreas and fetal kidney. Mouse embryonic expression studies show ubiquitous expression of the murine *Mipol1* gene during E10.5 to E13.5 stages. The 10th intron of the *MIPOL1* gene is reported to be disrupted in a t(2;14)(p23.3q13) patient affected with tetramelic mirror image polydactyly and left inguinal hernia (Kondoh et al., 2002).

5.3.2. The *PITX2* gene

The *PITX2* gene (paired-like homeodomain transcription factor 2), mapping at 4q25, is a member of the bicoid-like homeobox family of transcription factors. The gene, about 20 kb in size, has seven exons that encode four alternative transcripts (PITX2 A,B,C,D) (Cox et al., 2002). Each isoform is made by alternative splicing and differential use of three promoters. Two internal promoters reside in the third and fourth introns respectively. The protein sequences encoded by all isoforms share in common the homeodomain (encoded by exon 6) and carboxyl terminal domains (encoded by exon 7). The carboxyl terminal region contains the aristaless domain necessary for mediating protein-protein interactions. All isoforms contain dissimilar amino terminal domains, which are speculated to be required for the interaction of the PITX2 protein with other transcription factors. The PITX2B protein forms heterodimers with either PITX2A or PITX2C. One isoform (PITX2D), that has a truncated homeodomain and complete carboxyl terminal domain, negatively regulates PITX2 transcriptional activity by forming heterodimers with one of the three other isoforms and perhaps prevents interaction of the heterodimer with other transcription factors. The PITX2 protein also forms heterodimers with other transcription factors such as PIT1 to regulate promoters such as those of the prolactin (Amendt, et al., 1998), procollagen lysyl hydroxylase (Hjalt, et al., 2001) and distal-less homeobox 2 (Green, et al., 2001) genes. Studies predominantly in chick and mouse have suggested that the PITX2 isoforms act independently or in overlapping domains with different doses to pattern brain, tooth, heart, lung and gut development (Yu et al., 2001; Schweickert et al., 2000; Mucchielli, et al., 1997). Recent studies have also suggested other roles of the *PITX2* gene isoforms in hematopoietic stem cell differentiation (Degar et al., 2001), acute lymphocytic leukemia (Arakawa et al., 1998) and actin-myosin reorganization (Wei et al., 2002).

The *PITX2* gene was initially identified as the gene for Axenfeld-Rieger syndrome type 1 based on the finding of mutations predominantly within the homeodomain (Semina et al., 1996). Other studies have also shown the *PITX2* gene mutated in patients with Peter anomaly (Doward et al., 1999), iris hypoplasia and iridogoniodysgenesis (Kulak et al., 1998). Axenfield-Rieger syndrome type 1 (ARS1) is an autosomal dominant disorder, in which patients are affected with hypoplasia of the midface, teeth and a range of ocular defects. Other defects of the limb, heart, pituitary, genital and umbilical skin are also prevalent to varying extents among ARS1 patients.

About 50% of the patients develop glaucoma by age 20. In general, there is extensive clinical heterogeneity of the ARS1 phenotypic spectrum.

The *PITX2* gene is proposed to function downstream of the sonic hedgehog and nodal pathway in the left-right axis determination (Roessler and Muenke, 2001). Null alleles of the *Pitx2* gene were created in mice by several groups (Lin et al., 1999; Lu et al., 1999; Kitamura et al., 1999). Heterozygous and homozygous null *Pitx2* mice showed defects in the morphogenesis of the heart, mandibular and maxillary facial prominences, tooth, eye and pituitary; and in the establishment of asymmetry. In fact, many features of *Pitx2-/-* and *Pitx2+/-* mice are in common with those in Axenfeld-Rieger syndrome type 1 patients. However, no acallosal defects were seen among these mice since the concepti died before E15. A visible corpus callosum can be seen between E16 and E18 in mice. A hypomorphic allele of the *Pitx2* gene was also created with homozygous and heterozygous mice that were born at term (Gage et al., 1999). These mice again had features similar to the null *Pitx2* mice and Axenfeld-Rieger syndrome type 1 patients. However, no reports of acallosal defects were noted. It is possible that the difference in phenotype observed between mouse and human is a result of different roles of *PITX2* in these two species. The murine *Pitx2* gene is expressed in many tissues including the brain, heart, pituitary, mandibular and maxillary regions, eye, umbilicus and in the midline ventral domains of the early developing forebrain (Muccielli et al., 1996). The expression pattern of murine *Pitx2* suggests possible influences upon the midline-laterality determination by feedback control mechanisms with other midline genes such as sonic hedgehog, no tail and floating head.

5.4. METHODS AND MATERIALS

5.4.1. *Analysis of genomic sequence for genes at translocation breakpoint junctions*

Annotated genomic sequence was retrieved from Ensembl (www.ensembl.org), the National Centre for Bioinformatics Institute MapViewer (www.ncbi.nlm.nih.gov/genome/guide/human) to identify genes and to estimate physical map distances. BLAST searches were done on selected genomic BAC clone sequences to precisely map and estimate physical distances at the translocation breakpoint junction.

5.4.2. *High resolution breakpoint mapping*

Six chromosome 14q13-q21 and 14 chromosome 4q25 specific markers were selected from Ensembl and the WIGR STS content YAC maps for mapping using methods described in Chapter Two. An additional 12 amplimers were selected from Matsumoto et al. (1997) for more precise mapping at 14q13. Two markers within the *MIPOL1* gene, namely MPDK1 and MPDK2, and four markers 3' of the *PITX2* gene, namely R380DK1, R380DK2, R380DK3 and R380DK4 were designed from the *MIPOL1* cDNA sequence and the genomic sequence of BAC clone R380D23 (GI:18042428), using primer premiere 3. Three aliquots of each DOP-PCR amplified derivative chromosome were pooled before using for PCR analysis. Standard 20 µl PCR reactions were performed using 2 µl of pooled DOP-PCR pre-amplified flow sorted chromosomes of case 10. The PCR conditions and cycles are listed in Appendices 2, 13 and 14.

5.4.3. *Analysis of candidate genes for micro-rearrangement using flow sorted chromosomes*

Amplimers for the *NKX2.1*, *NKX2.9*, *SLC25A21*, *PITX2*, *PRO0971*, *T2BP* and *ANK2* genes were reported previously (UniSTS, www.ncbi.nlm.nih.gov) or designed with

Primer Premere 3 using cDNA sequences of these genes reported in Genbank. Three aliquots of each DOP-PCR amplified derivative chromosomes from case 10 were pooled before being used for standard 20 µl PCR analysis as outlined in Chapter Two. The PCR conditions and cycles used for each marker are reported in Appendix 15.

5.4.4. *Physical mapping of the PITX2 gene*

CEPH YAC clones - 910A2, 959H6, and 932A9 that contained markers binned between D4S193 and D4S1611 were selected from the WIGR. CEPH YAC clones were plated on AHC or YPD plates and grown 3-4 days at 30°C. Four to five single YAC colonies were chosen, pooled and grown in 5 mls of YPD medium overnight, shaking at 30°C. Mixed cultures were used in Whole Cell PCR as described in Chapter Three. PX3 amplimers, which flank exon six of the *PITX2* gene, were used for mapping onto YAC clones. The PCR conditions and cycles of PX3 are reported in Appendix 15.

5.4.5. *Analysis of the PITX2 gene for micro-rearrangement by densitometric analysis*

HindIII digested lymphoblast genomic DNA from four normal subjects and the proband were electrophoresed, subjected to Southern transfer onto Hybond N^+ filters (Amersham Pharmacia Biotech), and hybridized with ExpressHyb solution as specified by ClonTech. Co-hybridizations were done with a single copy 599 bp control probe and either a 432 bp PITX2 5' or 499 bp PITX2 3' probe. PCR primers for these probes were designed with Primer Premiere 3. A 5' PITX2 probe (PP1) was made with the amplimers (f-gagtccgtgctcctgctc/r-ctggacgtcagcagagattc). A 3' PITX2 probe (PP2) was made with the amplimers (f-ggtcgtgggcactaaagaaa/r-aaacatcattgcatccacca). A control probe (PC1) designed from BAC clone R340L03 was made with the amplimers (f-gcatggtgtcccttatttgc/r-tggctgagggctgaagttat). Each 20 µl PCR reaction contained 0.2 mM dNTP, 50 ng of each forward and reverse primer, 1.5 mM $MgCl_2$, 1X PCR buffer (Perkin Elmer), 50 ng human genomic DNA and 0.2 units of Taq polymerase (Sigma). The PCR cycle was denaturation at 94°C for 3 mins, 35 cycles of 94°C- 30 sec and 57°C - 30 sec, followed by a final extension of 72°C for 5 minutes. PCR products were electrophoresed on 1% agarose gels stained in ethidium bromide, excised and purified using the GenElute™ Gel Extraction Kit (Sigma). 50 ng of all probes were radiolabeled with ^{32}P-α-dCTP (10 mCi/ml) using the REDIprime Kit (Amersham Pharmacia Biotech). Washed blots were wrapped in Saran wrap and exposed in phosphorimager cassettes for up to 5 days. The phosphorimager screens were scanned with a Typhoon 8600 variable mode imager (Molecular Dynamics) and densitometric analysis was performed with the ImageQuant version 5.2 software (Molecular Dynamics). The ratio of signal intensity (counts per minute) was calculated in each lane between the control and PITX2 5' or 3' bands. All results were within the range of linearity of the instrument.

5.5. RESULTS

5.5.1. Brief clinical description

A description of case 10 is provided in Chapter Two. Briefly, this patient had a t(4;14)(q25q13) chromosome and was affected with agenesis of the corpus callosum, microcephaly, low set ears, flat face, bilateral epicanthal folds, feeding problems and developmental delay with no signs of ocular anomalies at the age of 7 months. The patient died during early childhood.

5.5.2. Fine breakpoint junction mapping

The translocation breakpoint junction was defined between the markers D4S2449 and Q9H1 on chromosome 4 using flow sorted chromosome analysis. Both markers resided within an estimated physical distance of 47 kb. The marker Q9H1 mapped at the 5' end of the *PRO0971* gene. Both D4S2449 and Q9H1 resided on BAC clone R477G18 (GI: 17136167). On chromosome 14, the translocation breakpoint junction was initially defined between the markers AFM200ZH4 and D14S306, an estimated 750 kb interval (See Chapter Two). However, by using a panel of amplimers reported by Matsumoto et al. (1997), the breakpoint junction was defined more precisely between the markers B319SP6 and B368T7 that resided on BAC clone R158A24 (GI:9967624). No deletions were found for any of the chromosomes 4 and 14 markers tested, except that the marker R380DK4, which mapped 25 kb distal to the 3' end of the PITX2 gene, was deleted (Figure 5-1).

5.5.3. Analysis of candidate genes for micro-rearrangements

Analysis of the breakpoint junction on chromosome 14q13 showed that the markers B368T7 and B319SP6 mapped within the third and fifth introns respectively of a recently identified gene known as *MIPOL1* (Kondoh et al., 2001) (Figure 5-2). The translocation breakpoint junction of case 10 at 14q13 was therefore between the third and fifth intron of *MIPOL1*. In order to precisely define the distal limits of the breakpoint junction in *MIPOL1*, physical mapping was done with an additional two markers, MPDK1 and MPDK2, within the fourth and fifth exons. Subsequent mapping of case 10 with these markers, showed that the breakpoint junction was between B368T7 and MPDK1, which is in intron 3 and in the 5' untranslated (UTR) interval of the *MIPOL1* gene. An examination of the genomic sequence at this breakpoint junction at 14q13 revealed that the 5' end of the *SLC25A21* gene (GI:13449278) was the closest gene, mapping within 71 kb proximal to the breakpoint. The *SLC25A21* gene (about 133 kb in size) encodes an

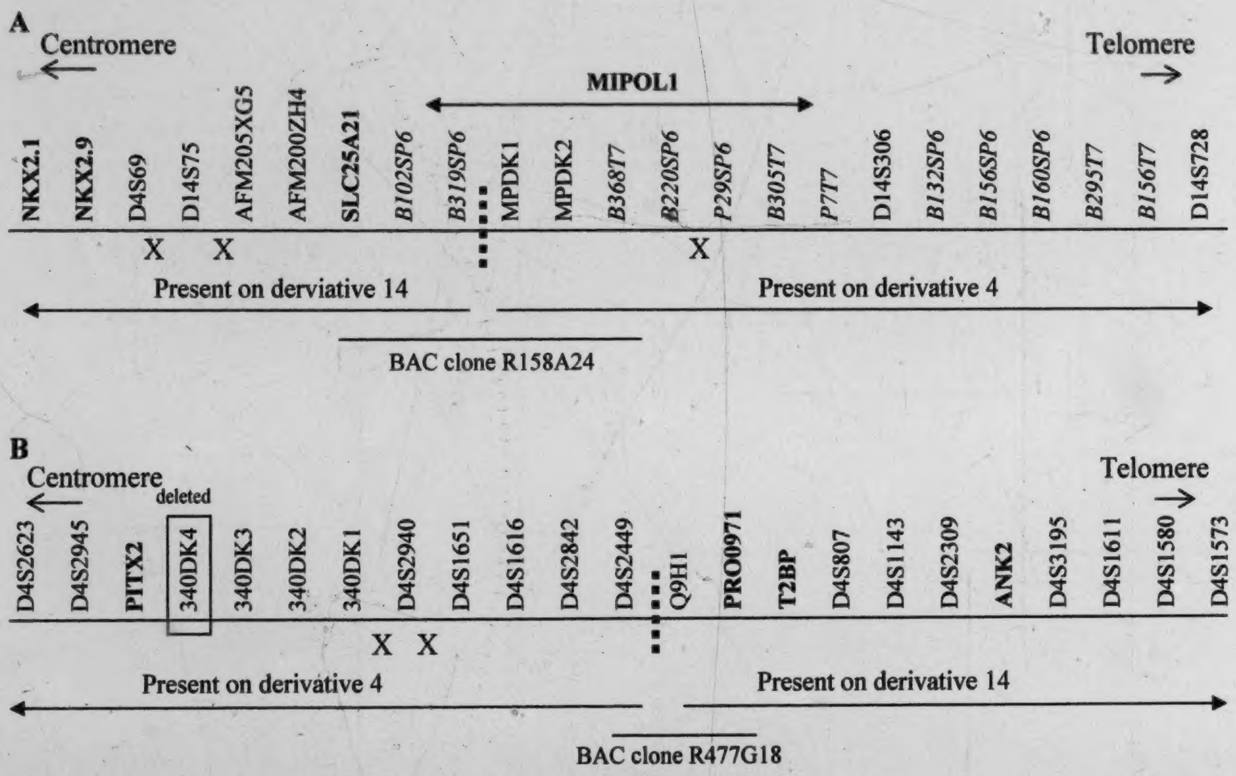

A
Centromere
Telomere
MIPOL1
NKX2.1
NKX2.9
D4S69
D14S75
AFM205XG5
AFM200ZH4
SLC25A21
B102SP6
B319SP6
MPDK1
MPDK2
B368T7
B220SP6
P29SP6
B305T7
P7T7
D14S306
B132SP6
B156SP6
B160SP6
B295T7
B156T7
D14S728
X
X
X
Present on derviative 14
Present on derivative 4
BAC clone R158A24
B
Centromere
deleted
Telomere
D4S2623
D4S2945
PITX2
340DK4
340DK3
340DK2
340DK1
D4S2940
D4S1651
D4S1616
D4S2842
D4S2449
Q9H1
PRO0971
T2BP
D4S807
D4S1143
D4S2309
ANK2
D4S3195
D4S1611
D4S1580
D4S1573
X
X
Present on derivative 4
Present on derivative 14
BAC clone R477G18

Figure 5-1: Breakpoint junction mapping of case 10 using flow sorted chromosomes. A) chromosome 14 analysis. B) chromosome 4 analysis. Genes are in bold; markers taken from Matsumoto et al., 1997 are in italics; stippled lines demarcate the breakpoint junction; X=other translocation cases (see text). The positions of markers are not drawn to scale. Of note, the markers B319SP6 and B368T7 are incorrectly reported in the reverse order by Kondoh et al., 2002.

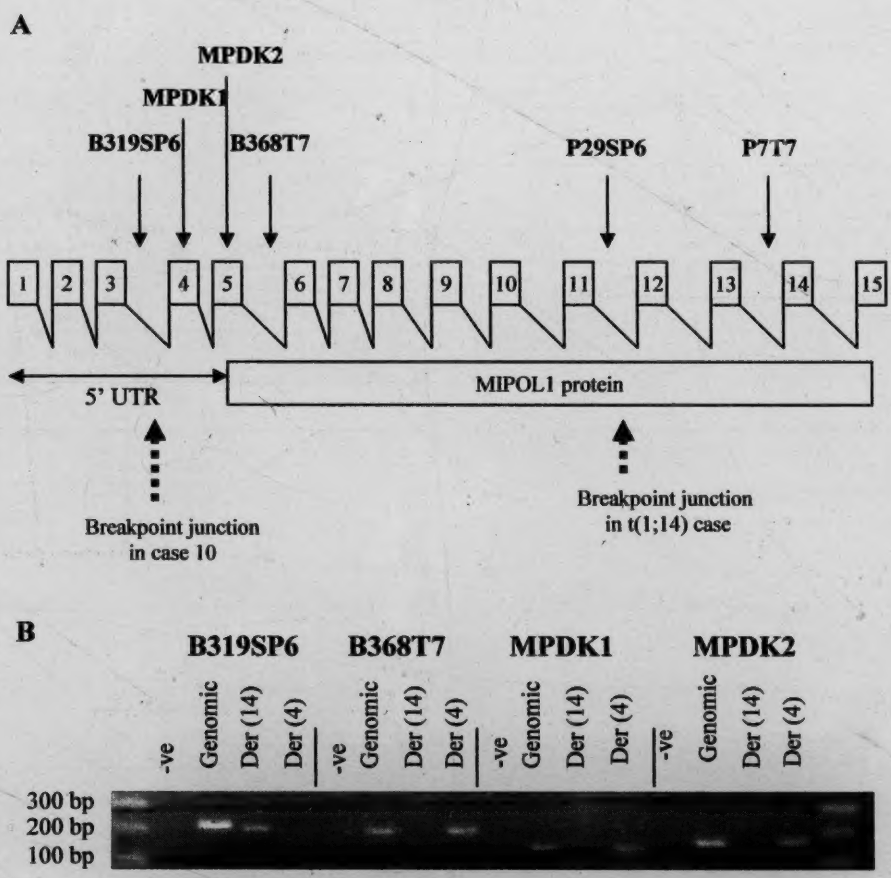

A
MPDK2
MPDK1
B319SP6
B368T7
P29SP6
P7T7
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
5' UTR
MIPOL1 protein
Breakpoint junction in case 10
Breakpoint junction in t(1;14) case
B
B319SP6
B368T7
MPDK1
MPDK2
-ve
Genomic
Der (14)
Der (4)
300 bp
200 bp
100 bp

Figure 5-2: Analysis of the *MIPOL1* gene in case 10 for rearrangements. A) schematic of the *MIPOL1* gene with markers used for mapping shown above. Not drawn to scale. Stippled arrows show translocation breakpoint junctions of case 10 and t(1;14) case. See Kondoh et al., 2001 for the latter case. Of note, the markers B319SP6 and B368T7 are incorrectly reported in the reverse order by Kondoh et al., 2001. B) flow sorted chromosome analysis with selected markers within the *MIPOL1* gene.

oxodicarboxylate transport protein, as part of a family of genes involved in mitochondrial metabolism (Fiermonte et al., 2001). *NKX2.1* (GI:4507714), about 3.8 kb in size and *NKX2.9* (GI:7657378), about 2 kb in size, physically map about 366 and 302 kb respectively, proximal to the breakpoint junction at 14q13. A homozygous knock-out of the murine *Nkx2.1* gene showed concepti with partial midline septal fusion defects of the central nervous system with absent structures including the mamillary bodies and pituitary (Kimura et al., 1995). The murine *Nkx2.9* has prominent expression in domains of the developing neural tube and ventral brain and is possibly involved in the patterning of a subset of neurons along the neuroaxis, in particular those of the ventral domains of the developing brain (Pabst et al., 1998). No murine knock-out model is yet reported for *Nkx2.9*. Using a panel of amplimers designed from the 5' and 3' ends of adjacent genes *SLC25A21*, *NKX2.1* and *NKX2.9* of 14q13, no deletions or rearrangements were detected (Figures 5-3).

Analysis of the genomic sequence of 4q25 demonstrated that the 5' end of the *PROO971* gene (about 38 kb in size), a putative gene with unknown function but partial similarity to kinases (GI:8923996), is the closest gene mapping within 47 kb distal to the breakpoint junction. The next closest gene mapping about 91 kb distal to the breakpoint junction at 4q25, is *T2BP* (GI: 17436598), about 9.8 kb in size, has an unknown function and identity. The *ANK2* gene, another excellent candidate, mapped at least 950kb distal to the breakpoint junction. The *ANK2* gene, about 333 kb in size with three transcript isoforms, belongs to a family of integral membrane proteins involved in diverse roles such as cell motility, activation, proliferation, and cell-cell contact. Studies in rat and *Drosophila* previously showed that the *ANK2* orthologue had predominant expression in

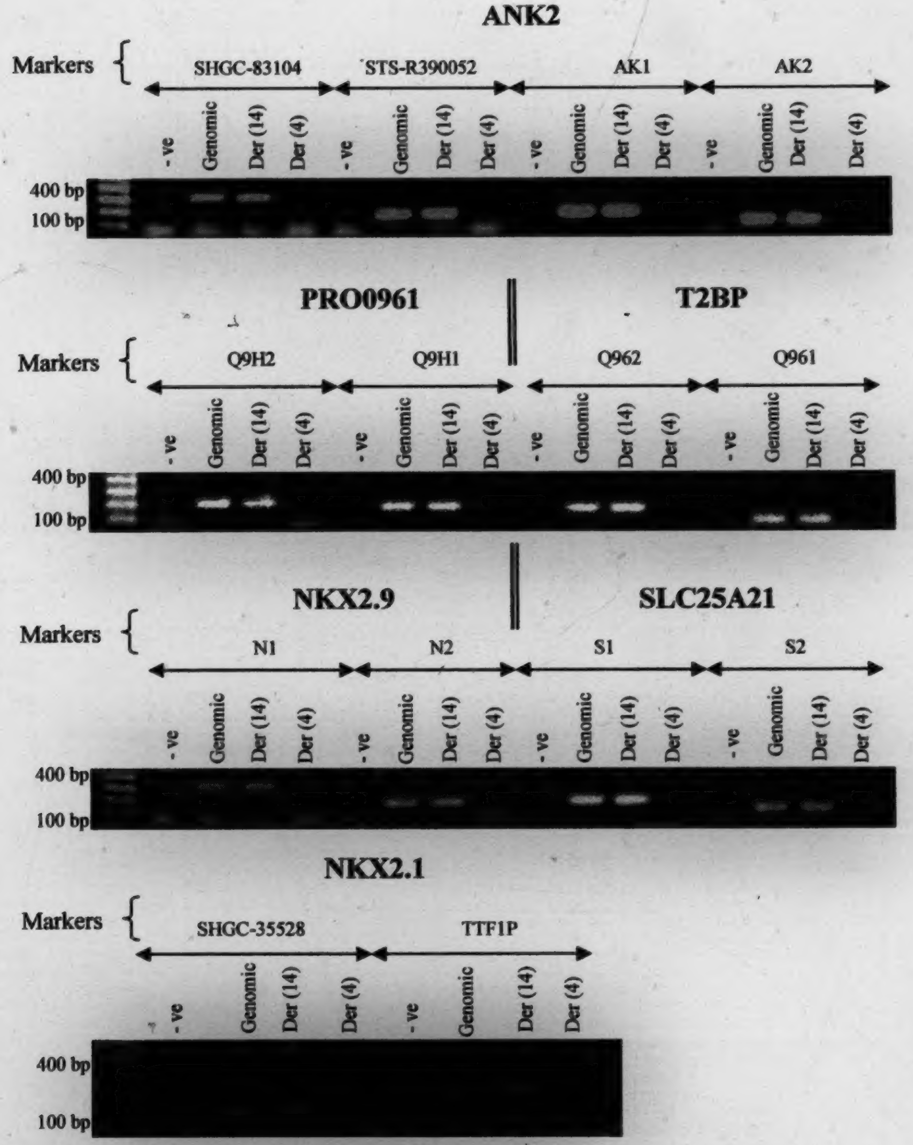

ANK2
Markers
SHGC-83104
STS-R390052
AK1
AK2
- ve
Genomic
Der (14)
Der (4)
400 bp
100 bp
PRO0961
T2BP
Q9H2
Q9H1
Q962
Q961
NKX2.9
SLC25A21
N1
N2
S1
S2
NKX2.1
SHGC-35528
TTF1P

Figure 5-3: Flow sorted chromosome analysis with markers mapping within a panel of candidate genes flanking the translocation breakpoint junctions on chromosomes 14q13 and 4q25 in case 10. Der=derivative chromosome

embryonic, neonatal and adult brain, particularly in areas rich in axons and dendrites (Bouley et al., 2000; Chan et al., 1993). The Ank2 protein is selectively targeted to premyelinated axons, and is possibly involved in neuronal migration. Using amplimers designed within nearby genes *PRO0971*, *T2BP* and *ANK2*, no deletions or rearrangements were detected among these genes (Figure 5-3). *PITX2* was examined for potential rearrangements based on previous case reports of its association with patients with hypoplasia or agenesis of the corpus callosum (Makita et al., 1995; Kulharya et al., 1995). Physical mapping onto CEPH YAC clones in this dissertation had estimated the *PITX2* gene at a maximum of 5.6 Mb proximal to the breakpoint junction at 4q25. However, with the availability of draft human genomic sequence, the *PITX2* gene was estimated to map about 1.7 Mb proximal to the translocation breakpoint junction. Using amplimers mapping to specific exons of the *PITX2* gene, a deletion was found from exons 6 to 7 on flow sorted derivative 4 and 14 chromosomes of case 10, an estimated interval of at least 3.9 kb. Exons 1 to 5 were still present on the derivative 4 chromosome (Figure 5-4 a,b). Densitometric analysis was used to confirm these findings within *PITX2*. Specifically, the 5' end of *PITX2* was present in two copies and the 3' end was present in one copy (Figure 5-4 a,c). The R380DK4 marker was deleted, indicating a deletion of about 25 kb at the 3' end of the *PITX2* gene, in conjunction with the deletion of exons six and seven, suggest as large as a 50 kb microdeletion within and distal to the *PITX2* gene.

5.6. DISCUSSION

Case 10 was examined in detail since this patient was affected with agenesis of the corpus callosum, which was evident in six of the deletion cases reported in Chapter Two.

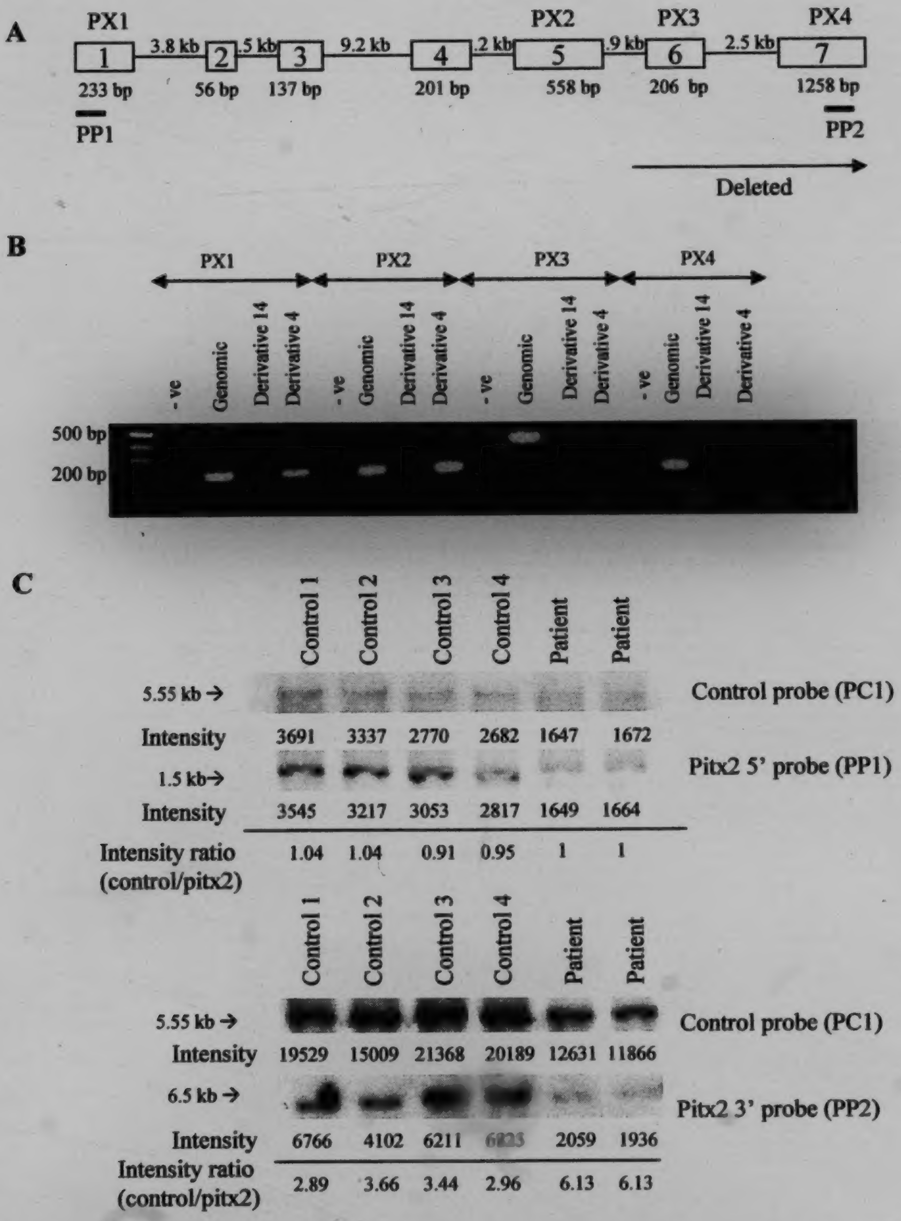

A
PX1
PX2
PX3
PX4
1
2
3
4
5
6
7
3.8 kb
.5 kb
9.2 kb
.2 kb
.9 kb
2.5 kb
233 bp
56 bp
137 bp
201 bp
558 bp
206 bp
1258 bp
PP1
PP2
Deleted
B
PX1
PX2
PX3
PX4
- ve
Genomic
Derivative 14
Derivative 4
- ve
Genomic
Derivative 14
Derivative 4
- ve
Genomic
Derivative 14
Derivative 4
- ve
Genomic
Derivative 14
Derivative 4
500 bp
200 bp
C
Control 1
Control 2
Control 3
Control 4
Patient
Patient
5.55 kb →
Control probe (PC1)
Intensity
3691 3337 2770 2682 1647 1672
1.5 kb →
Pitx2 5' probe (PP1)
Intensity
3545 3217 3053 2817 1649 1664
Intensity ratio (control/pitx2)
1.04 1.04 0.91 0.95 1 1
Control 1
Control 2
Control 3
Control 4
Patient
Patient
5.55 kb →
Control probe (PC1)
Intensity
19529 15009 21368 20189 12631 11866
6.5 kb →
Pitx2 3' probe (PP2)
Intensity
6766 4102 6211 6823 2059 1936
Intensity ratio (control/pitx2)
2.89 3.66 3.44 2.96 6.13 6.13

Figure 5-4: Analysis of the *PITX2* gene for rearrangements in case 10. A) schematic of the *PITX2* gene. Positions of probes (PP1, PP2) used for densitometry are shown. Not drawn to scale. B) flow sorted chromosome analysis with selected primers mapping within the *PITX2* gene. C) densitometric analysis of a control and 5' or 3' *PITX2* probes.

The chromosome 14 translocation breakpoint junction of this patient, between the markers AFM200ZH4 and D14S306, mapped over 100 kb distal to the initially defined holoprosencephaly locus between the markers D14S49 and AFM205XG5. Furthermore, this patient's translocation breakpoint junction mapped over 2.2 Mb distal to the redefined holoprosencephaly locus between the markers D14S49 and D14S1014. Nevertheless, the search for one or more genes on chromosomes 14 and 4 as the cause of agenesis of the corpus callosum was pursued.

The translocation breakpoint junctions were defined on chromosomes 4q25 and 14q13 by mapping on flow sorted aberrant chromosomes. Analysis of candidate genes within a 2 Mb interval at the 14q13 and 4q25 breakpoint junctions showed no deletions or rearrangements in *PRO0916*, *T2BP*, *ANK2*, *SLC25A21*, *NKX2.1* and *NKX2.9* thus excluding them as potentially contributing to the phenotype in this patient. However, the possibility of long range position effects exerted on the transcriptional regulation of these genes cannot be excluded. The translocation breakpoint junction was found within the third intron of the *MIPOL1* gene at 14q13 (Kondoh et al., 2002), which disrupted the 5' UTR of the transcript. On chromosome 4q25, a translocation breakpoint junction mapped between D4S2449 and Q9H1. In addition, a microdeletion as large as 50 kb was found within and distal to the 3' end of the *PITX2* gene. As part of the microdeletion on 4q25, a deletion of at least a 3.9 kb, including exons 6 and 7 of *PITX2* was found. A translocation breakpoint in conjunction with a neighboring deletion implied that the interval at 4q25 was possibly inverted in a complex micro-rearrangement. There are two possible models for this. In the first model, an unbalanced translocation breakpoint within *PITX2* fused with *MIOPL1*, followed by an inversion of the interval between the

3' end of *PITX2* and the markers D4S2449 and Q9H1. In the second model, a translocation breakpoint between the markers D4S2449 and Q9H1 fused with *MIOPL1*, followed by an inversion of a proximal 1.7 Mb interval that disrupted and deleted part of the 3' end of *PITX2*. This rearrangement was unlikely to be detected in the proband's karyotype due to insufficient metaphase GTG banding resolution.

Most remarkable is the finding of a patient affected with tetramelic mirror image polydactyly and no central nervous system anomalies, and a disruption of the 11th intron of the *MIPOL1* gene (Kondoh et al., 2002). Since the carboxyl terminal region of the MIPOL1 protein has been postulated to contain functional domains such as coiled-coil domains (Kondoh et al.,2002), the translocation breakpoint junction of the tetramelic mirror image polydactyly patient likely affected the coding potential and function of the disrupted protein. Case 10, who had no tetramelic mirror image polydactyly, was demonstrated to have a translocation breakpoint junction in the 5' UTR of the *MIPOL1* gene. Given this finding, the disrupted *MIPOL1* gene in case 10 is likely to be functional since the coding region of the protein is not affected. A promoter within the third intron of the disrupted *MIPOL1* gene at 14q13 or even a promoter on 4q25 could serve to regulate normal transcription of this disrupted gene, explaining why case 10 did not have a phenotype in common with the other patient (Kondoh et al., 2002). Two other patients (case 8, case 9) described in Chapter Two also had translocation breakpoints at 14q13 that mapped within 300 kb of the breakpoint of case 10. These other patients had no features in common with each other or with case 10.

The deletion of exons 6 and 7 of the *PITX2* gene in this patient is possibly a cause of the agenesis of the corpus callosum. The deletion in the 3' end of *PITX2* is predicted to delete the entire carboxyl terminal domain and homeodomain of all four *PITX2* transcript isoforms thereby creating a null allele. No other genes are found within the microdeletion, estimated to be as large as 50 kb extending 3' of the *PITX2* gene. Hence, haploinsufficiency of the *PITX2* gene cannot be excluded as the cause of the phenotype. Functional studies performed on a selection of Axenfeld-Rieger syndrome type 1 disease causing mutations show these mutations reduced slightly or completely abrogate the transactivation or binding ability of the PITX2 protein (Priston et al., 2001; Kozlowski et al., 2000). One *PITX2* homeodomain mutation caused slight cytoplasmic mislocalization of the protein (Priston et al., 2001). Based on these findings, Axenfeld-Rieger syndrome type 1 patients were proposed to have only one copy, or hyperactive copy, in addition to a functional *PITX2* to cause an abnormal phenotype. None of the ARS1 patients tested for *PITX2* mutations had MRI or CT cranial imaging reported. Thus it is unknown whether these patients had acallosal defects as observed in case 10. For case 10, it is unknown whether the child had glaucoma or any dental anomalies manifested within the ARS1 phenotype spectrum. Other patients with *PITX2* mutations may also be non-penetrant for the ocular and dental anomalies that are mainly associated with the ARS1 phenotypic spectrum, and would not be initially selected for *PITX2* mutation screening.

Patients with 4q25 translocations and deletions have provided evidence of the phenotypic spectrum which includes those seen in case 10. Three molecularly defined translocation cases involving the 4q25 band have been reported previously (Datson et al., 1996; Folmen et al., 1998). In all three cases, the translocation breakpoint junctions mapped within 100 to 200 kb distal to *PITX2* and all patients were affected with the Axenfeld-Reiger syndrome type 1 spectrum including ocular, dental and maxillary hypoplasia. Among these cases, *PITX2* was not reported to be rearranged or deleted. The translocation breakpoint junctions among these cases may have exerted position effects on the regulation of *PITX2*. A complex rearrangement involving chromosomes 3,4,10,17 was reported in a patient with ARS1 features, however the breakpoints were undefined (Ogilvie et al., 1998). Another uncharacterized t(1;4)(q23.1;q25) case was reported in a patient with ARS1 ocular and umbilical defects (Makita et al., 1995). This patient also had central nervous system midline defects including hypoplasia of the corpus callosum and agenesis of the anterior commissure. Clinical heterogeneity is observed in the 4q25 deletion cases (Kulharya et al., 1995; Schinzel et al., 1997). Of the 12 cases reported, only five had typical ARS1 features. Three cases had complete deletions of the 4q25 band with no ARS1 phenotypes but a spectrum of CNS anomalies (Kulharya et al., 1995). Most 4q25 deletion cases have clinical features in common with case 10, who has only one functional copy of *PITX2*. Therefore, a locus for agenesis of the corpus callosum resides at 4q25 and may possibly involve the *PITX2* gene.

5.7. SECTION B – ANALYSIS OF FAMILIAL t(9;14) CASES 11 AND 12 (Cell lines L6874 and SMOM)

5.8. METHODS AND MATERIALS

5.8.1. *Analysis of genomic sequence for genes at translocation breakpoint junctions*

Annotation of the genomic sequence was retrieved from Ensembl (www.ensembl.org), the National Centre for Bioinformatics Institute MapViewer (www.ncbi.nlm.nih.gov/genome/guide/human), and the HPE8 critical region described in Chapters Three and Four, to identify genes and to estimate physical map distances. BLAST searches were done on selected genomic BAC clone sequences to precisely map and estimate physical distances at the translocation breakpoint junction.

5.8.2. Fine breakpoint mapping

Sixteen amplimers, namely, I2DK, I3DK, I4DK, I5DK, I6DK, I7DK, I8DK, SDK1, SDK2, SDK5, AEX2, 1075M22SL, 1075M22SR, 1078I14SR, 66M11SL and 66M11SL, were designed from the sequences of BAC clones R1075M22 (GI:16215182), R1078I14 (GI:14272173) and R66M11 (GI:14268348), using primer premiere 3, to precisely map the chromosome 14q breakpoint junction. Amplimers K1DK, K2DK, K3DK, K4DK, SDK6 and SDK7 mapping within intron three, exons three and four of the KIAA0391 gene (GI:4240256), were designed from the cDNA and genomic sequence retrieved from Genbank, using primer premiere 3. To define the translocation breakpoint junction on chromosome 9q, sixteen chromosome 9 specific polymorphic and non-polymorphic markers (Research Genetics) were selected from the WIGR STS content YAC maps (www-genome.wi.mit.edu). The order of markers used in mapping was ascertained from the WIGR YAC physical maps and any available human chromosomes 9 and 14 genomic sequences (GenBank, www.ncbi.nlm.nih.gov). DOP-PCR amplification of aberrant chromosomes for case 11 and 12 is described in Chapter Two. Three aliquots of each DOP-PCR amplified derivative chromosomes were pooled before being used for PCR analysis. The DNA template for each standard 20 µl PCR reaction included either 40 ng genomic DNA, or 2 µl of DOP-PCR amplified flow sorted aberrant chromosomes. The PCR conditions and cycles are listed in Appendix 16.

5.9. RESULTS

5.9.1. Brief clinical description

Cases 11 and 12 are described in Chapter Two. Case 11, mother of case 12, has schizophrenia and mild mental delay. Case 12, the daughter of case 11, has schizophreniform psychosis, developmental delay, severe mental delay. Both family members had a t(9;14)(q34;q13) chromosome.

5.9.2. Fine breakpoint mapping on chromosome 14

In Chapter Two, the translocation breakpoint junctions of cases 11 and 12 were described between the markers D14S730 and D14S70, an estimated 683 Kb. In order to define more precisely the breakpoint junction on chromosome 14q13, flow sorted chromosome analyses were performed with an additional sixteen amplimers designed from BAC clone or cDNA sequences. Physical mapping with these markers had demonstrated the translocation breakpoint junctions of both cases 11 and 12 between 66M11SR and SDK1, an estimated 7.9 kb (Figure 5-5). Fifteen of the 16 markers mapped within the *NPAS3* gene, newly identified and characterized (Chapters Three and Four). The translocation breakpoint junction was identified within the third intron of this gene since both 66M11SR and SDK1 markers mapped within this interval (Figures 5-5, 5-6). The translocation breakpoint junctions of both cases disrupted both alternative transcripts of *NPAS3* (Figure 5-6). The rearrangement in this gene was confirmed by metaphase FISH analysis in collaboration with Dr. Muir. Figure 5-7 shows the use of BAC clones R372C16 and R1075M22 as probes that mapped respectively to the 5' and 3' regions of *NPAS3*, with metaphase FISH on the lymphoblast cell line of case 12. The same finding was evident in the mother.

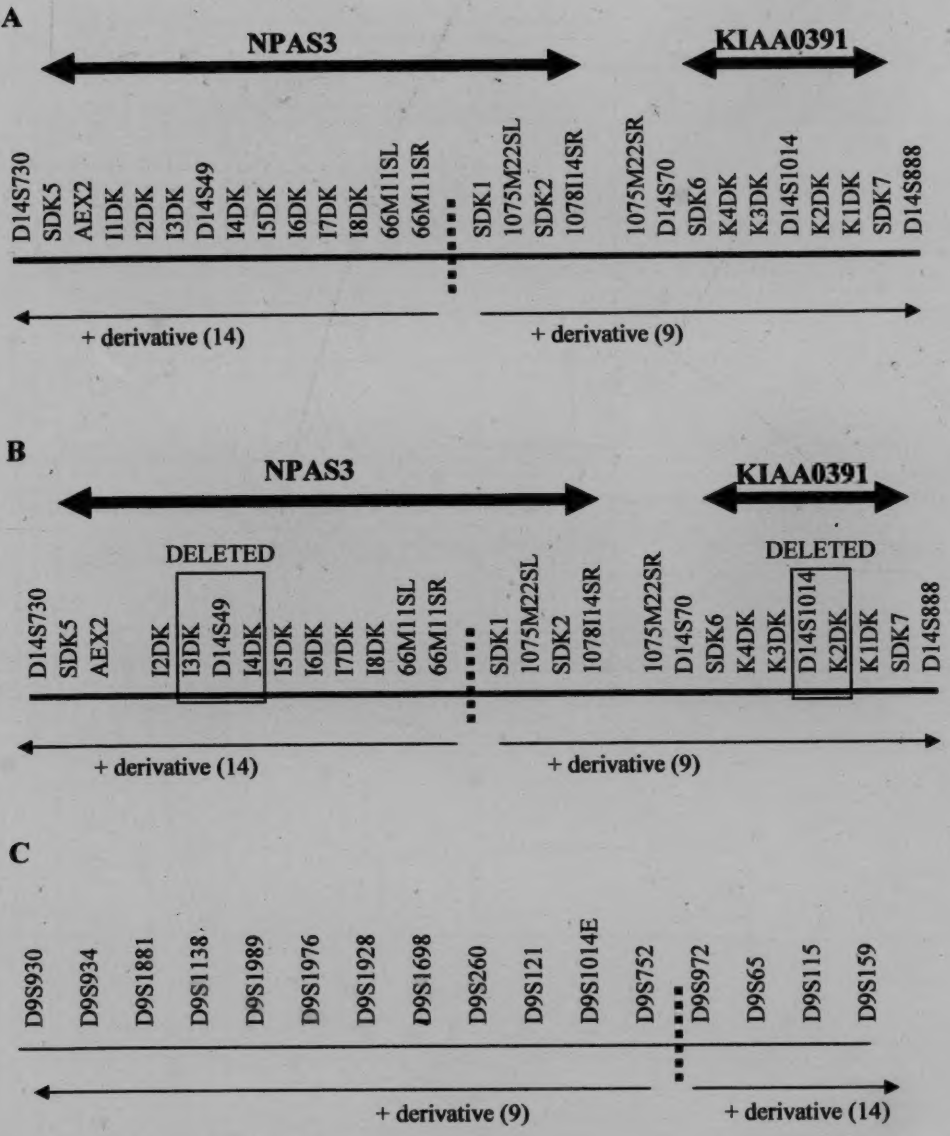

A
NPAS3
KIAA0391
D14S730
SDK5
AEX2
I1DK
I2DK
I3DK
D14S49
I4DK
I5DK
I6DK
I7DK
I8DK
66M11SL
66M11SR
SDK1
1075M22SL
SDK2
1078I14SR
1075M22SR
D14S70
SDK6
K4DK
K3DK
D14S1014
K2DK
K1DK
SDK7
D14S888
+ derivative (14)
+ derivative (9)
B
NPAS3
KIAA0391
DELETED
DELETED
D14S730
SDK5
AEX2
I2DK
I3DK
D14S49
I4DK
I5DK
I6DK
I7DK
I8DK
66M11SL
66M11SR
SDK1
1075M22SL
SDK2
1078I14SR
1075M22SR
D14S70
SDK6
K4DK
K3DK
D14S1014
K2DK
K1DK
SDK7
D14S888
+ derivative (14)
+ derivative (9)
C
D9S930
D9S934
D9S1881
D9S1138
D9S1989
D9S1976
D9S1928
D9S1698
D9S260
D9S121
D9S1014E
D9S752
D9S972
D9S65
D9S115
D9S159
+ derivative (9)
+ derivative (14)

Figure 5-5: Breakpoint junction analysis of chromosomes 14 and 9 of cases 11 and 12. A) chromosome 14q13 analysis of case 11 (mother). B) chromosome 14q13 analyses of case 12 (daughter). C) chromosome 9q34 analysis of mother and daughter. Stippled lines = translocation breakpoint junction. Positions of markers are not drawn to scale.

The daughter (case 12), was found to have additional micro-rearrangements at 14q13, not detected in the mother. Deletions of three of nine markers within a maximum estimated 94 kb interval were found within the second intron of *NPAS3* (Figure 5-5). Initial analysis had identified a deletion of the D14S1014 marker that maps within the third intron of the *KIAA0391* gene (Figure 5-5). Further mapping within this gene showed a deletion of an additional one out of four markers within the third intron of *KIAA0391* (GI:4240256), an estimated maximum deletion of 22 kb (Figure 5-5). The *KIAA0391* gene has unknown function and identity and maps about 1.2 Mb distal to *NPAS3*. This finding suggests a complex rearrangement at 14q13 involving one or more inversions in conjunction with a translocation event. Fiber-FISH analyses with a panel of BAC clones (R1075M22, R372C16 and R588D7) however failed to demonstrate any complex rearrangements in case 12 (Benjamin Pickard, personal communication).

5.9.3. Fine breakpoint mapping of cases 11 and 12 on chromosome 9

On chromosome 9q34, the translocation breakpoint junction was defined between D9S752 and D9S972, an estimated 100 kb, in both cases 11 and 12 (Figure 5-5). No deletions or complex rearrangements on chromosome 9 were detected with sixteen markers tested and no genes were disrupted within this interval. The closest genes mapping within 1 Mb to the breakpoint junction at 9q34 were three genes with unknown functions (KIAA0169 (GI: 22046117), KIAA1848 (GI: 14017912), LOC204994 (GI: 22043509)), five hypothetical genes (LOC255259 (GI: 22046125), LOC169627 (GI: 18572511), LOC206943 (GI: 20538549), LOC169656 (GI: 20538553), LOC138519 (GI: 17469193)) and four known genes (LSFR2 (GI: 22046119), CRAT (GI: 21618330), PPP2R4 (GI: 10880986), AD-003 (GI: 7661527)).

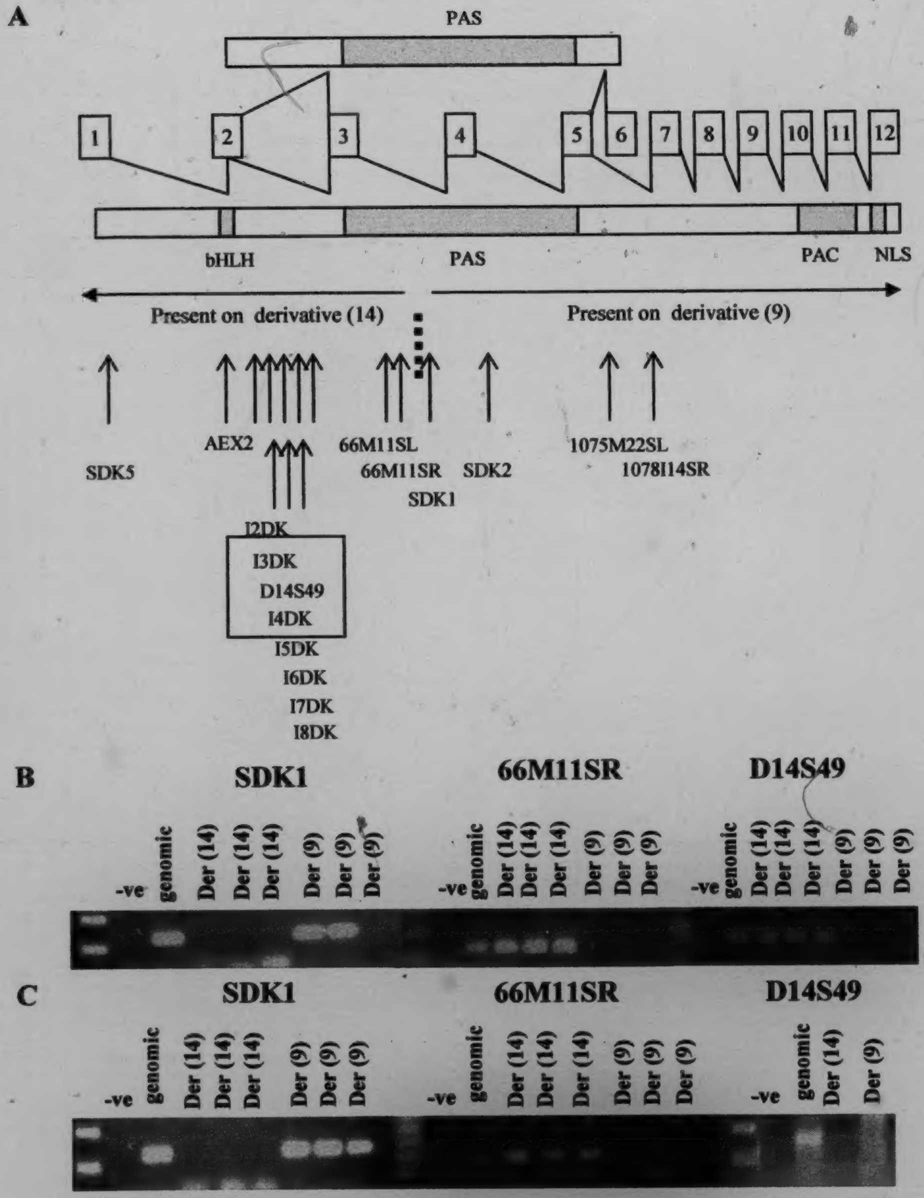

A
PAS
1
2
3
4
5
6
7
8
9
10
11
12
bHLH
PAS
PAC
NLS
Present on derivative (14)
Present on derivative (9)
SDK5
AEX2
66M11SL
66M11SR
SDK2
SDK1
1075M22SL
1078I14SR
I2DK
I3DK
D14S49
I4DK
I5DK
I6DK
I7DK
I8DK
B
SDK1
-ve
genomic
Der (14)
Der (14)
Der (14)
Der (9)
Der (9)
Der (9)
66M11SR
-ve
genomic
Der (14)
Der (14)
Der (14)
Der (9)
Der (9)
Der (9)
D14S49
-ve
genomic
Der (14)
Der (14)
Der (14)
Der (9)
Der (9)
Der (9)
C
SDK1
-ve
genomic
Der (14)
Der (14)
Der (14)
Der (9)
Der (9)
Der (9)
66M11SR
-ve
genomic
Der (14)
Der (14)
Der (14)
Der (9)
Der (9)
Der (9)
D14S49
-ve
genomic
Der (14)
Der (9)

Figure 5-6: Search for rearrangement in the *NPAS3* gene in the t(9;14) family. A) schematic of *NPAS3* showing positions of markers used in mapping. Stippled lines=translocation breakpoint junction. The functional motifs/domains of the encoded proteins are shown. Not drawn to scale. B) flow sorted chromosome analysis of case 11 (mother) with selected markers within *NPAS3*. C) flow sorted chromosome analyses of case 12 (daughter) with selected markers within *NPAS3*. The der(14) and der(9) templates were separate or pools of three DOP-PCR amplified der(14) or der(9) flow sorted chromosomes.

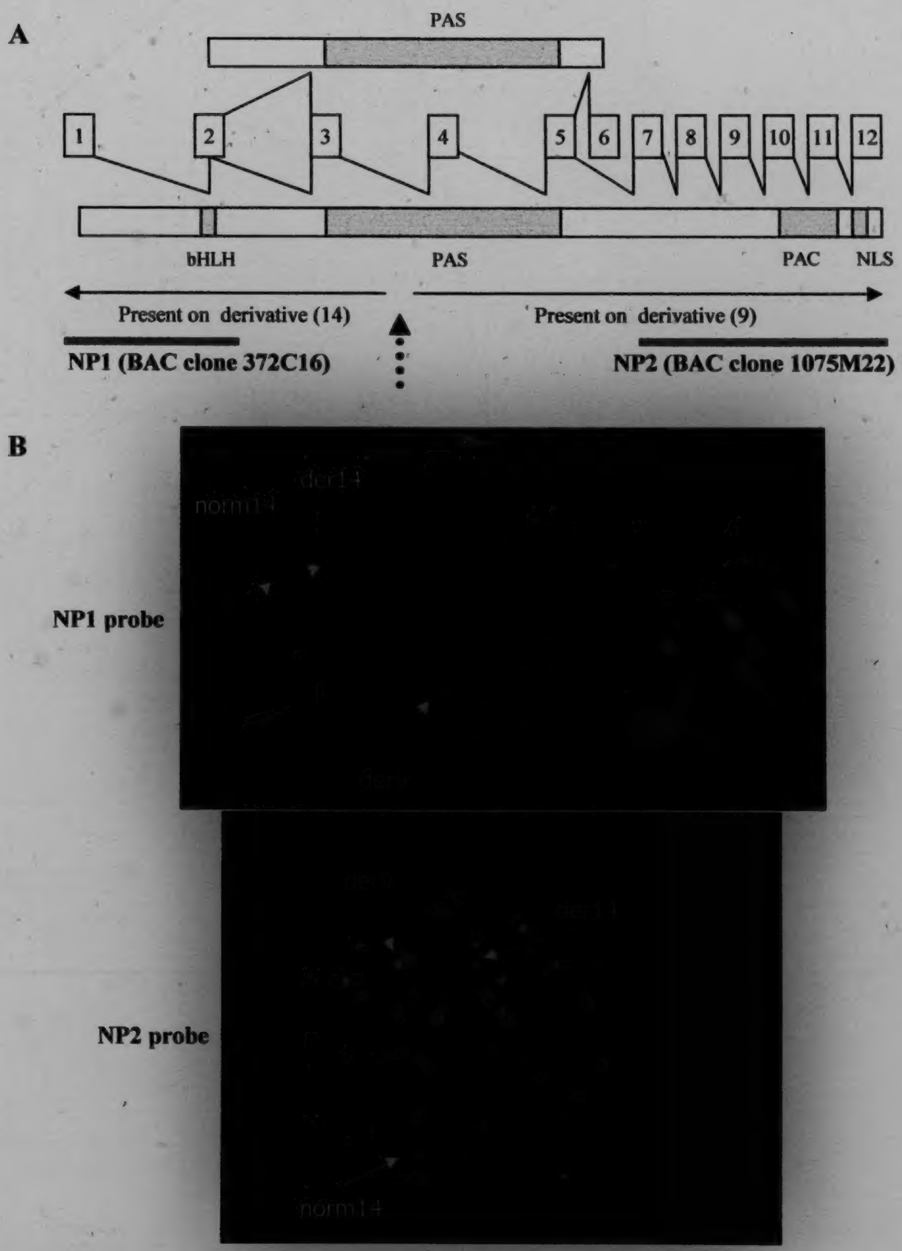

A
PAS
1
2
3
4
5
6
7
8
9
10
11
12
bHLH
PAS
PAC
NLS
Present on derivative (14)
Present on derivative (9)
NP1 (BAC clone 372C16)
NP2 (BAC clone 1075M22)
B
NP1 probe
der14
norm14
der9
NP2 probe
der9
der14
norm14

Figure 5-7: Identification of the rearrangement in the *NPAS3* gene in the t(9;14) family. A) schematic of *NPAS3* showing positions of probes (NP1 and NP2) used for metaphase FISH analysis. The functional motifs/domains of the protein isoforms are shown. Not drawn to scale. Stippled arrow=translocation breakpoint junction. B) metaphase FISH analysis on a lymphoblast cell line from case 12, with BAC clone probes specific to the 5' (NP1) and 3' (NP2) ends of *NPAS3*. These figures were provided by Dr. Benjamin Picard and Dr. Muir (University of Edinburgh, UK).

5.10. DISCUSION

Cases 11 and 12 were examined in detail since both mother and daughter were found to have translocation breakpoint junctions between the markers D14S730 and D14S70 as reported in Chapter Two. This interval included part of the centromeric boundary of the defined holoprosencephaly locus between the markers D14S49 and D14S1014. Both mother and daughter were affected with mental illness which was not reported among any of the deletion cases mapped in Chapter Two, in addition to other proximal chromosome 14q13 deletion cases reported in the literature. However, this may possibly be due to the fact that the subjects were too young to manifest psychiatric problems. In addition, the possibility of co-morbid occurrence of mental illness with the chromosome aberration in this family cannot be excluded. The translocation breakpoint junctions of both cases were examined to determine whether one or more genes within the defined HPE8 locus could be the cause of the phenotype observed.

Both daughter and mother had a translocation breakpoint junction in the third intron of *NPAS3* which disrupted the coding potential of both alternative transcripts: the first 124 amino acids of the amino terminus of the larger 901 amino acid isoform, and the first 52 amino acids of the amino terminus of the putative smaller 153 amino acid isoform. The bHLH (basic helix-loop-helix) domain of the larger protein was disrupted, preventing the protein from binding to DNA. For both protein isoforms, the PAS domain that is required for dimerization is disrupted. The PAC motif and bipartite nuclear localization signal within the carboxyl terminus of the larger protein isoform remains intact. An abnormal NPAS3 protein may contribute to the cause of the phenotype observed in this family. Thus far, only the *SIM2* gene of the bHLH-PAS gene family of transcription factors, mapping on human chromosome 22, has been associated with behavioral problems as seen among Down syndrome patients (Chrast et al., 2000). None of the members of this transcription factor family, however, has been associated with schizophrenia. The findings of this chapter suggest that a bHLH-PAS gene may be associated with schizophrenia.

Based on the mapping data, the daughter (case 11) appeared to have a more complex rearrangement than the mother (case 12) on chromosome 14. This is typical for familial translocation cases since translocation chromosomes frequently rearrange during meiosis due to improper synapsing that leads to unequal recombination (reviewed by Shaffer and Lupski, 2000). The additional deletion can explain the more severe phenotype in the daughter. Microdeletions were found within the second intron of *NPAS3* and third intron of *KIAA0391* mapping slightly over 1 Mb distal to *NPAS3*. The identity and function of the *KIAA0391* gene is unknown. In addition, it is unknown how these microdeletions affect these genes. The finding of additional rearrangements at 9q34 in case 12 had suggested the possibility of one or more inversions in conjunction to a translocation breakpoint junction at 14q13. This rationale however was not supported since Fiber-FISH with three chromosome 14q13 BAC clones did not demonstrate any complex rearrangements at 14q13 (Benjamin Picard, personal communication). The fact that the daughter inherited a genetic background from a father who had bipolar disorder can also explain the possibility of more severe phenotypes manifested in her.

There is some evidence, although not strong, for a locus for schizophrenia at chromosome 14q13 (reviewed in Chapter Two). Since the *NPAS3* gene was suggested as an excellent candidate for holoprosencephaly in Chapter Four, with the data of this chapter, this gene still cannot be excluded as a holoprosencephaly candidate. This stems from previous studies that reported families affected with the holoprosencephaly spectrum in addition to schizophrenia or other mental illnesses (Roach, et al., 1975; Hercig et al., 1994). Interestingly, these subjects had normal or mild facial signs (Roach et al., 1975; Hercig et al., 1994). Unfortunately, because of cognitive and physical impairment, it is extremely difficult to assess mental disorders in those affected with holoprosencephaly. Holoprosencephaly patients with behavioral problems, noted thus far tend to have excessive crying, irritability, self-stimulation, and self-abusive behaviors. Many of the patients die at an early age before any proper psychiatric assessment can be performed. It has not been possible to obtain cranial MRI or CT-scans of either the mother or daughter. Thus, it is unknown whether these patients have central nervous system anomalies within the holoprosencephaly spectrum.

On chromosome 9q34, the breakpoint junction was defined between D9S952 and D9S972, an estimated 100 kb in both cases 11 and 12. No genes were found disrupted within this interval. However the possibility of position mediated effects on the transcriptional expression of neighboring genes cannot be excluded. There was evidence for weak association of the *NMDAR1* gene at chromosome 9q34 with schizophrenia in a South African Bantu-speaking tribe (Riley et al., 1997). The NMDAR1 receptor functions in the glutaminergic pathway and was demonstrated in mice by means of a hypomorphic allele to result in schizophrenia-like behavior that could be treated with haloperidol (Mohn et al., 1999). However, this gene maps several Mb from the chromosome 9q34 breakpoint junction of cases 11 and 12. There is one case report of a 28 year old man with mental retardation, schizophrenia, short stature, short webbed neck, dysmorphic face, and mild anomalies of the fingers, who had a del(9)(q32q34) (Park et al., 1991). The deletion interval in this patient is yet to be characterized.

5.11. SUMMARY OF FINDINGS

SECTION A

1) A complex rearrangement involving a disruption of the *MIPOL1* gene at 14q13 and a disruption and/or microdeletion of at least 3.9 kb in the 3' end of the *PITX2* gene was found in a t(4;14) patient affected with acallosal defects and mild craniofacial findings.

2) Haploinsufficiency of the *PITX2* gene at 4q25 was proposed as a possible cause of the phenotype seen in the t(4;14) patient.

SECTION B

1) A translocation breakpoint junction was found within the third intron of the *NPAS3* gene in a family with a t(9;14) chromosome and affected with schizophrenia.

2) The disruption of the *NPAS3* gene was proposed to possibly contribute to schizophrenia, however this gene remains an excellent candidate for holoprosencephaly since previous studies have demonstrated holoprosencephaly families with schizophrenia or other mental illnesses.

CHAPTER SIX

GENERAL DISCUSSION AND FUTURE DIRECTIONS

6.1. THE HYPOTHESIS AND SUMMARY OF CONCLUSIONS

The general hypothesis presented at the beginning of this dissertation was the following: *"The proximal region of human chromosome 14, namely, 14q13, contains one or more genes that when mutated cause the holoprosencephaly spectrum of phenotypes, including severe alobar holoprosencephaly to mild holoprosencephaly features".*

This hypothesis arose from the premise that a small panel of patients with sporadic chromosome 14 aberrations had phenotypes within severe to mild forms of holoprosencephaly. The fact that the cytogenetic deleted intervals of classic holoprosencephaly cases were overlapping with those of patients with features within mild holoprosencephaly suggested that the spectrum of midline anomalies of the central nervous system was caused by a common deleted cytogenetic interval, which is 14q13. Haploinsufficiency of one or more genes at the 14q13 locus is therefore a proposed explanation for the molecular pathogenesis of the holoprosencephaly spectrum of phenotypes observed with patients with proximal chromosome 14q13 rearrangements. Linkage analysis on holoprosencephaly families is yet to be performed in order to define additional HPE loci other than HPE3. This information will be supportive to the finding of a locus for holoprosencephaly on chromosome 14q13. The shortcoming with the cases designated with features within mild holoprosencephaly is that the previous cases (cases 1,2,4,5,6) were reported as having mainly complete or partial agenesis of the corpus callosum. These cases (1,2,4,5,6) were designated mild holoprosencephaly spectrum cases since the proximal chromosome 14q deletion intervals overlapped with the classic holoprosencephaly forebrain defect cases. Most likely the acallosal defect cases represented a mild variant of the classical holoprosencephaly forebrain anomalies as seen in patients with syntelencephaly and lobar holoprosencephaly. Furthermore, one subject, that is, case 6, had acallosal defects in addition to white matter heterotopia and an underdeveloped pituitary gland as seen in syntelencephaly and lobar holoprosencephaly. The clinical descriptions obtained from reporting physicians were possibly incomplete, and in some instances the cranial images were not interpreted by experts. To support this notion, when there is partial or complete agenesis of the corpus callosum, the development of the septum pellucidum is partial or abolished, the lateral ventricles become enlarged and the third ventricle becomes displaced. The description of the cases in Chapter Two does not include these other anomalies that are typically noted with acallosal defects, thereby supporting the possible incompleteness of the clinical information. Furthermore, details on partial agenesis of the corpus callosum were unknown as to whether the anterior or central corpus callosum was absent as seen in syntelencephaly, semilobar and lobar holoprosencephaly.

The data described in this dissertation were based on testing several sub-hypotheses.

1) *A locus for holoprosencephaly (HPE8) exists at human chromosome 14q13.* Chapter Two defined a locus for holoprosencephaly by the use of overlapping deletions from two patients who had alobar holoprosencephaly or mild holoprosencephaly spectrum. An estimated 4.2 Mb interval was defined between the markers D14S49 and AFM205XG5. By using the distal boundary of a deletion patient exhibiting no holoprosencephaly features (case 13), the distal portion of the holoprosencephaly locus was excluded, leaving the defined interval between the markers D14S49 and D14S1014, an estimated 2 Mb interval. Of the mapped deletion cases, one was classic alobar

holoprosencephaly, five were within the mild holoprosencephaly spectrum and three did not have holoprosencephaly. The molecular characterization of the deletion breakpoints suggested that both classic holoprosencephaly and mild HPE spectrum shared the minimal critical region between the markers D14S49 and D14S1014, and therefore could be caused by the same molecular pathogenesis. The HPE8 locus was found to result in extreme clinically heterogeneity. There was no evidence for parent of origin effects and no correlations were determined between sex or ethnicity and the phenotype severity. These findings are consistent with the findings of other well studied and characterized holoprosencephaly loci on chromosomes 2 (HPE2),7 (HPE3),13 (HPE5),18 (HPE4) and 21 (HPE1).

2) *The HPE8 locus contains one or more candidate genes for holoprosencephaly*

Chapter Three described the building of a 3 Mb YAC contig, which was the prerequisite for the construction of a 2 Mb BAC contig. These contigs eventually led to the construction of a transcript map comprising 34 genes, by extensive annotation of available draft human genomic sequence from the Human Genome Project and Celera Genomics. Five potential holoprosencephaly genes, namely *SNX6*, *NPAS3*, *C14ORF11*, *C14ORF10* and *C14ORF19*, were initially proposed based on the fact that these genes are expressed in human fetal brain at 21 to 30 weeks gestation. A gene for holoprosencephaly should express between three and seven weeks of gestation, since the central nervous system and craniofacial architecture are established during this period. A holoprosencephaly gene would be expected to maintain expression in selected central nervous system regions up to 30 weeks of gestation. Typically, during 21 to 30 weeks of gestation, the three fissures form. Two genes, namely *C14ORF10* and *C14ORF19* were excluded for further analysis as HPE candidates. *C14ORF10* was more likely a candidate for idiopathic basal ganglia calcification linked to D14S1014, that mapped less than 500 kb distal to this gene. For the other excluded gene, further analysis was not possible since an incomplete cDNA sequence was obtained. However the gene was related to immunoglobulin binding proteins that function in the immune system and therefore was an unlikely candidate. Out of 34 possible genes annotated from the 2 Mb of genomic sequence between the markers D14S49 and D14S1014, three genes were selected as candidates for holoprosencephaly. The 2.2 Mb interval between D14S1014 and AFM205XG5, that was excluded by patients who did not demonstrate holoprosencephaly features, was reassessed for potential holoprosencephaly candidate genes since these patients could possibly show non-penetrance. An analysis of a panel of 10 genes mapping to this interval showed only one (FLJ23250 fis) that was expressed in human fetal brain at 21 to 33 weeks gestation. This gene was not a promising holoprosencephaly candidate since it was similar to transposases.

3) *The candidate genes of HPE8 have properties similar to other known holoprosencephaly genes*

The functional characterization of three holoprosencephaly candidate genes was described in Chapter Four using resources from human and mouse. Studies in humans included the identification of the cDNA, genomic structure, *in silico* expression profile, adult multi tissue expression profile and protein cellular localization. These findings have provided some insight into the biology of these genes. Studies in mouse have provided the temporal embryonic developmental expression profile of the murine orthologues and the spatial expression during neurulation at 11 dpc.

All three genes varied in size and genomic structure. The *NPAS3* gene was found to have unusually large introns varying from 274 to 116 kb in size. By *in silico* expression profiling, all three candidate genes were shown to be expressed in a multitude of fetal and adult tissues, in addition to several cancer cell lines. Developmental genes are often implicated in the cause of cancer. For instance, the *SHH* and *PTCH* genes are mutated in patients with holoprosencephaly and also in several acquired forms of skin cancer (Hahn et al., 1999). All three candidate genes are expressed in human adult tissues. Specifically, *SNX6* and *C14ORF11* were expressed ubiquitously. The larger transcript of *NPAS3* was expressed only in human adult brain. The smaller transcript of this gene was expressed only in human fetal brain. All three genes, including the alternative transcripts, expressed as early as 21 to 33 weeks gestation in the human fetal brain. Testing for expression in the developing human brain prior to 21 weeks of gestation was restricted due to sample unavailability. For the seven genes found mutated thus far among sporadic and familial HPE cases, namely *SHH, ZIC2, TGIF, PTCH, TDGF1, SIX3* and *DKK1*, the detailed expression patterns in human adult tissues are mostly unknown and therefore require further investigation. However, the *TDGF1* gene is known to be expressed in the adult colon using a limited panel of tissues (Minchiotti et al., 2002).

In this dissertation, the proteins of two of the three candidate genes were localized to the cytoplasm of COS1 and human transformed skin fibroblast cell lines. Other studies have reported the localization of the SNX6 protein in both the cytoplasm and nucleus of HeLa cells (Teasdale et al., 2001; Ishibashi et al., 2001). The NPAS3 protein was localized to the nucleus of COS1 and transformed human skin fibroblast cell lines, supporting its role as a transcription factor, since the beta domain is known to function in DNA binding in other family members of this gene (Crews, 1998). The C14ORF11 and NPAS3 proteins could have multiple cellular localizations depending on the cell types and spatio-temporal expression. The *SNX6* gene is involved in the trafficking of Nodal receptors and ser/thr receptor kinases, and the *C14ORF11* gene is possibly an accessory protein. The cellular localization and possible roles of these three candidate genes are consistent with the diverse roles ascertained thus far for known holoprosencephaly genes.

The C57BL/6J mouse strain was used as a model to study the developmental expression profile of the murine orthologues for the three holoprosencephaly candidate genes. The murine orthologues of all three holoprosencephaly candidate genes were expressed as early as gastrulation and continued until birth. All three genes were expressed in the developing central nervous system, neural tube and craniofacial structures at 11 dpc. The *C14ORF11* and *SNX6* genes were also expressed in the developing limbs and viscera. *SHH, PTCH, ZIC2, TDGF1, DKK1* and *SIX3* have been demonstrated to express as early as gastrulation in the developing mouse (Roelink et al., 1995; Platt et al., 1997; Nagai et al., 1997; Oliver et al., 1995; Kawakami et al., 1996; Ding et al., 1998). Studies on the earlier expression profiling, that is before 14 dpc, of the murine *Tgif* gene are warranted. All seven known HPE genes also show similar spatial expression profiles with respect to the proposed three candidate genes. In summary, the three holoprosencephaly candidate genes at 14q13 have human and murine functional properties similar to those of other known HPE genes.

4) *Translocation cases are important in excluding or including genes and disease phenotypes that are associated with a chromosome*

None of the translocation cases characterized in Chapter Two shared common phenotypes with the phenotypic spectrum observed in the proximal chromosome 14q deletion cases, except for a t(4;14)(q25;q13) case affected with agenesis of the corpus callosum. Chapter Five described the detailed characterization of this t(4;14) case to determine whether the gene that caused agenesis of the corpus callosum was on chromosome 4q25 or chromosome 14q13. This patient was found to have a translocation breakpoint junction in the 5' UTR of the *MIPOL1* gene at 14q13 which was also reported to be disrupted in the 3' end in another patient with tetramelic mirror image polydactyly and no brain anomalies (Kondoh et al., 2002). Since the translocation breakpoint junction was in the 5' UTR of the *MIPOL1* gene, this study proposed that the coding potential of this gene, that is responsible for limb anomalies that were not seen in the t(4;14) patient, was still functional due to the activation of a cryptic promoter distal to the breakpoint junction at 14q13 or a promoter on 4q25. Further mapping on this patient also revealed a complex rearrangement on 4q25 with an approximate 50 kb deletion that resulted in deletion of 3.4 kb of the 3' end of the *PITX2* gene. A second rearrangement was found about 1.7 Mb distal to the 3' end of the *PITX2* gene. A translocation and an inversion event were proposed to occur at 4q25 in this patient. The deletion within the 3' of the *PITX2* gene was predicted to produce a null allele as a result of the loss of the bicoid-like homeodomain and carboxyl terminal domain. With this finding, and from a previous study that associated the *PITX2* gene with agenesis of the corpus callosum (Makita et al., 1995), this dissertation proposed that the agenesis of the corpus callosum phenotype observed in the t(4;14) patient may be due to the rearrangement in *PITX2* at 4q25. This finding was important, as it showed a link between the midline and laterality developmental pathways in humans. Furthermore, it has suggested that patients diagnosed with Axenfeld-Reiger type I syndrome and glaucoma might be investigated by cranial imaging to discern any structural central nervous system anomalies.

Chapter Five also described a mother and daughter with a t(9;14)(q34;q13) chromosome and affected with schizophrenia. The breakpoint junction was within the centromeric boundary of the HPE8 locus. Detailed mapping using flow sorted chromosome analysis revealed that the breakpoint junction was within the third intron of *NPAS3* disrupting the coding ability of both alternative transcripts of this gene. Both mother and daughter were found to have a common translocation breakpoint junction on chromosome 9q34, however the daughter also had a 94 kb microdeletion within the second intron of *NPAS3* and a second 22 kb microdeletion within the third intron of the *KIAA0391* gene mapping 1.2 Mb distal to *NPAS3*. The functional relevance of these microdeletions is unknown. The daughter's phenotype, more severe than that of the mother, can be explained by these additional microdeletions and the fact that she had inherited a genetic background from her father who was also mentally ill. Since no genes were found to be disrupted at 9q34, the *NPAS3* gene at 14q13 was proposed to possibly contribute to the schizophrenia phenotype observed in this family. Other genes of this bHLH-PAS gene family such as *NPAS2* and *SIM2*, are associated respectively with circadian problems or behavioral problems as seen in Down syndrome patients (Chrast et al., 2000). However, these findings did not exclude the *NPAS3* gene as a holoprosencephaly candidate since previous studies have demonstrated families with holoprosencephaly and normal faces that are affected with mental illness, including schizophrenia (Roach et al., 1975; Hercig et al., 1994). In summary, the three translocation cases mentioned above have led to proposing possible causative gene on chromosome 14 or on the other derivative chromosome.

6.2. PARALLELS BETWEEN THE CHROMOSOME 14q13 AND 13q32 HPE LOCI

Of the 11 other proposed HPE loci, HPE8 at 14q13 and HPE5 at 13q32 are most similar with respect to the clinical findings. The HPE5 locus was initially defined as part of the 13q- deletion syndrome, that later assisted in defining the holoprosencephaly critical region containing the *ZIC2* gene (Brown et al., 1995, 2001). No 13q32 deletion cases have yet been reported with single central maxillary incisors, as seen with the 14q13 deletion cases. The 13q32 deletion patients demonstrated alobar to lobar classic holoprosencephaly but only mild facial findings, such as intercanthal spacing problems, flat nose, clefting of the lip and/or palate or iris colobomas (Brown et al., 1995, 1998). The latter finding is also observed among the 14q13 deletion cases. Interestingly, holoprosencephaly patients with *ZIC2* mutations were found to have either normal faces or mild facial anomalies as seen among the 13q32 deletion cases (Brown et al., 1998, 2001). These patients demonstrated alobar, semilobar or lobar classic HPE forebrain anomalies. However, one patient with mild facial signs was found to have the middle interhemispheric variant of mild holoprosencephaly or syntelencephaly (Brown et al., 1998, 2001). The HPE8 and HPE5 loci likely contain many non-diagnostic HPE facial signs that are part of the estimated 20% that do not follow the "DeMyer face-brain relationship". A gene for holoprosencephaly at 14q13 may therefore have spatio temporal embryonic patterning roles similar to the *ZIC2* gene in order to manifest such similar clinical findings. It is unknown where, during human embryogenesis, haploinsufficiency of *ZIC2* can result in the holoprosencephaly phenotype. The *ZIC2* gene is known thus far to function in the odd paired and sonic hedgehog signaling pathways (Nagai et al., 1997; Koyabu et al., 2001; Mizugishi et al., 2001). It is also interesting that a murine *Zic2* hypomorphic allele resulted in behavioral anomalies consistent with schizophrenia (Ogura et al., 2001). This finding is most remarkable since Chapter Five described a family affected with schizophrenia and having a disruption within the *NPAS3* holoprosencephaly candidate gene. Whether the *NPAS3* gene is the causative gene for holoprosencephaly at HPE8 is yet to be established upon mutation analyses on a panel of holoprosencephaly patients. Both the murine *Zic2* (Nagai et al., 1997) and *Npas3* genes have similar developmental expression profiles in the manner that they are expressed as early as gastrulation and later in the developing central nervous system, spinal column and limbs. Furthermore, this dissertation showed the human *NPAS3* gene had restricted expression of one of the alternative transcripts in the adult brain. The murine *Zic2* gene also was reported to have restricted expression in adult brain (Nagai et al., 1997).

6.3. A CHROMOSOME 14q13 DELETION SYNDROME IS UNLIKELY THE CAUSE OF THE PHENOTYPE

The definition of a syndrome is important in the identification of etiologically distinct entities which inevitably will assist in the elucidation and comprehension of the molecular pathogenesis of such syndromes (Cohen, 1989c). Holoprosencephaly is an extremely complex disease, having challenged the principles of syndromology. This stems from the fact that holoprosencephaly has been associated with unknown syndromes and also with known syndromes including familial, chromosomal and environmentally-induced syndromes (Cohen, 1981). The proposal of a chromosome 14q13 deletion syndrome as the cause of the holoprosencephaly phenotype is based on the fact that no correlations were found between the deletion sizes of the mapped deletion cases, including those reported in the literature (Ramelli et al., 2000; Govaerts et al., 1996; Bruyere et al., 1996; Iwatani et al., 2000; Levin and Surana, 1991) and the severity of the holoprosencephaly spectrum of phenotypes. Taken together, haploinsufficiency of a single gene is likely responsible for the holoprosencephaly spectrum of phenotypes observed in the proximal chromosome 14q deletion cases. Variable expression of the holoprosencephaly spectrum of phenotypes at 14q13 is also demonstrated for the other proposed 11 loci on 11 other chromosomes. Variable expression of the phenotype can result from the fact that different genes are deleted, each patient has a different genetic background, the penetrance may vary or the gene-environment interaction during gestation may vary.

6.4. MODELS OF HOW HPE8 CANDIDATE GENES CAUSE A HOLOPROSENCEPHALY PHENOTYPE

The *NPAS3* gene was proposed as an excellent candidate for holoprosencephaly based on 1) expression in human fetal brain (21-30 weeks gestation), 2) expression as early as gastrulation in the C57BL/6J mouse embryos and 3) expression in the developing central nervous system of 11 dpc C57BL/6J mouse embryos. Hence this gene is transcriptionally active during the relevant embryonic periods for the genesis of the holoprosencephaly phenotype. This gene likely encodes a transcription factor since it was found localized in the nucleus and the beta domain is known to bind to DNA. The function of this transcription factor is unknown, however proteins with PAS domains are known to dimerize with each other (Crews and Fan, 1999). This transcription factor may function in toxin metabolism, circadian rhythmn, tracheal development, hypoxic response or even neurogenesis (Crews and Fan, 1999). The Single minded (*Sim*) orthologue of the *NPAS3* gene is a master regulator of midline development of the central nervous system and results in fused ventricles reminiscent of holoprosencephaly if mutant (Thomas et al., 1988; Nambu et al., 1991). The *Sim2* murine orthologue of the *Drosophila Sim* gene has been shown to possibly regulate the expression of *Shh* in the ventral midbrain and diencephalon (Epstein et al., 2000). Given this, the *NPAS3* holoprosencephaly candidate gene is proposed to function in the Sonic Hedgehog signaling pathway. A mutant NPAS3 protein could result in altered expression of the Sonic Hedghog ligand, thereby affecting sonic hedgehog signaling as predicted in patients with mutations in the *SHH* or *PTCH* genes.

The *C14ORF11* gene is also another excellent candidate for holoprosencephaly for the same three reasons mentioned above for the *NPAS3* gene. The encoded protein is not similar to any proteins or functional protein motifs in the database, except its murine ortholog with unknown function, thus it is very challenging to propose predicted functions. This dissertation had suggested that the gene encodes an accessory protein that is localized at least in the cytoplasm of COS1 or transformed human skin fibroblast cell lines. The developmental pathway in which this gene functions is unknown. For these reasons, a possible model of how this gene functions in the genesis of the holoprosencephaly phenotype cannot be proposed with presently available information.

The *SNX6* gene is a third gene excellent candidate for holoprosencephaly for the same three reasons mentioned above for the *NPAS3* gene. The work of other groups (Parks et al., 2001) has also suggested that the SNX6 protein interacts with the transforming growth factor receptor II, activin receptor type II B, and serine/threonine kinases that regulate these receptors, thereby resulting in the trafficking of these proteins to their appropriate cellular localization for the establishment of the nodal signaling pathway. A mutant SNX6 protein is therefore proposed to aberrantly traffic the nodal receptor and the associated Ser/Thr nodal receptor kinases to their appropriate cellular localization. Based on these previous findings and those reported in this thesis, the *SNX6* gene seems to be the most likely of the three proposed candidate genes for holoprosencephaly at chromosome 14q13. A non-functional SNX6 protein is predicted to alter nodal receptor signaling and/or regulation of the nodal receptor signaling by its Ser/Thr nodal receptor kinases, either by insufficient localization or mislocalization of the interacting proteins. The possibility of binding to the interacting nodal receptor proteins and sequestering them from functioning in a dominant negative manner, is another possibility associated with a non-functional SNX6 protein. Since the nodal receptor signaling pathway is critical in defining the axis of the developing embryo as early as gastrulation, aberrant signaling is predicted to be lethal. This finding is supported from *Snx1-/-Snx2-/-* double homozygous mice demonstrating arrest at midgestation with open anterior neural folds and underdeveloped rostral head formation (Schwarz et al., 2002). The SNX6 protein has been shown to form heterodimers with either SNX1 or SNX2 (Parks et al., 2001). Furthermore, SNX1 and SNX2 proteins also interact with a subset of nodal receptor kinases (Worby and Dixon, 2002), suggesting the SNX6 mutant phenotype may be similar to those of SNX1/SNX2 double mutants. Creation of a murine *Snx6* knock-out with or without *Snx1* and/or *Snx2* null backgrounds, and the screening a panel of holoprosencephaly patients will resolve whether this gene causes a holoprosencephaly phenotype if haploinsufficient.

6.5. SIGNIFICANCE OF THIS DISSERTATION

The findings of this dissertation have led to the definition of a 2 Mb minimal critical region for holoprosencephaly at chromosome 14q13 with three proposed candidate genes. Whether one or two of these three genes may, when mutated, cause the holoprosencephaly spectrum of phenotypes observed in patients with proximal chromosome 14q13 will be discerned upon mutation screening of a panel of holoprosencephaly patients. The finding of a holoprosencephaly gene obtained from the preliminary findings and proposals from this dissertation will be important in understanding the molecular pathways that contribute to the human embryogenesis, in particular the development of the prosencephalon. Furthermore, it will assist in understanding the molecular pathogenesis of the holoprosencephaly phenotype. Specifically, what are the molecular pathways ? What genes are the key players ? How do these genes interact to form the holoprosencephaly phenotype ? The holoprosencephaly spectrum of phenotypes is also challenging by the inability to discern whether a fetus or child is affected or unaffected with holoprosencephaly based on incomplete clinical information or variable expression of the holoprosencephaly spectrum. Given this, screening for mutations in a panel of genome wide holoprosencephaly genes, including those at HPE8, will resolve ambiguities in diagnosing the phenotype by molecular means, and subsequently assist in the management of the fetus or child.

6.6. FUTURE DIRECTIONS

The translocation breakpoint junctions of six cases are defined in this dissertation. The breakpoint junctions of three subjects, namely cases 8, 9, and 12, need more precise definition on chromosome 14 and on the other derivative chromosome to determine whether genes were disrupted or mapped close to the breakpoint junction. Based on the defined breakpoint junctions of these three cases, no genes were found disrupted. Furthermore, these subjects do not have any common features with the panel of deletion cases examined. This suggests that the other derivative chromosome may contain the causative gene(s) or gain of function of a gene on chromosome 14 due to aberrant gene expression could result in the phenotype.

This dissertation also reported no evidence for uniparental disomy of chromosomes 14 and 22, and proximal chromosome 14q13 microdeletion in a subject (case 16) with a Robertsonian t(14q22q) chromosome and clinical microforms within the holoprosencephaly spectrum. The search for the cause of the phenotype in this child requires additional studies. These include mapping with the use of flow sorted t(14q22q) chromosomes with a more dense panel of chromosomes 14 and 22 markers in search of microdeletions outside 14q13. Furthermore, mapping by FISH can ascertain any structural chromosomal rearrangements such as inversions or duplication on chromosomes 14 or 22 in this subject as the cause of the phenotype.

Analysis of a t(9;14) chromosome in a family described two microdeletions, specifically 94 kb in size within the second intron of *NPAS3,* and 22 kb in size within the third intron of *KIAA0391*, in a subject affected with schizophreniform psychosis and developmental (case 12). The genomic sequences deleted in this subject contain many predicted potential transcription factor binding sites. The functional relevance of these deleted regions requires further studies. In addition, the function of the *KIAA0391* gene needs to be elucidated. These studies will assist in understanding the more severe phenotype of this subject when compared with her mother (case 11). The molecular findings of this family also suggest that the *NPAS3* gene is an excellent candidate to screen for mutations in patients affected with schizophrenia.

The *C14ORF10* gene was isolated and partly characterized in this dissertation. This gene is similar to the gene family of calcium/calmodulin kinases, and maps within 500 kb of the D14S1014 marker that showed the highest linkage in patients affected with Fahr disease (Geschwind et al., 1999). Functional studies on the C14ORF10 protein are needed to demonstrate that it is a kinase. Since the major clinical hallmark of Fahr disease is aberrant calcium homeostasis in the central nervous system, the *C14ORF10* gene is an excellent candidate to screen for mutations in these patients.

The three candidate genes for holoprosencephaly at 14q13 (*SNX6*, *NPAS3*, *C14ORF11*) have been characterized to some extent in human and mouse. Since the genomic structures of these genes are known, mutation analyses will need to be performed on a panel of holoprosencephaly sporadic and familial cases by collaborative efforts. The finding of mutations in one or more of these genes among holoprosencephaly patients will be definitive in the identification of holoprosencephaly gene(s) at 14q13. The proteomic characterizations of these genes will ultimately provide insight into the features and functional properties of these proteins, and will be crucial to understanding the functional effects of mutations if found among holoprosencephaly patients. Such studies include DNA-protein binding assays for the NPAS3 protein, protein-protein interactions and phosphorylation site assays.

This dissertation reported the temporal expression profile of the murine orthologues of the three HPE candidate genes as early as gastrulation. The temporal expression profiling in post natal and adult tissues in the developing mouse is needed. More thorough spatial expression profiling is warranted for these genes since only 11 dpc C57BL/6J embryos were examined by whole mount *in situ* hybridization. Spatial expression profiles from days 6 to 18, in addition to specific post natal and adult tissues need to be obtained to understand the roles of these genes during development.

Finally, thorough developmental expression profiling studies can be supplemented by gene targeted disruption in mouse, in order to ascertain the phenotype and to make phenotype correlations with the holoprosencephaly spectrum. The use of embryonic stem cell transgenesis is an example of a method to create mouse models. Specifically, murine embryonic stem (ES) cells can be transfected with a gene targeting vector that contains part of the targeted gene sequence and a *LacZ-Neomycin* cassette. Gene targeted ES cells, after selection with G418, are microinjected into donor blastocysts, then the chimeric blastocysts are implanted into pseudopregnant mice. Southern blotting, PCR, quantitative RT-PCR, and/or western blotting can be used to genotype possible chimeric pups to confirm that these mice harbor a disrupted gene. Chimeric carriers are bred to obtain heterozygous and homozygous pups, and phenotypes such are structural defects are documented, if any. Since the human *NPAS3* gene was found disrupted in a family with mental illness, the murine targeted *Npas3*, can be accessed for behavioral problems, if any, as performed with the murine *Zic2* known-down model for holoprosencephaly (Ogura et al., 2001). If no pups survive to term pregnancy, embryos at various developmental stages will need to be harvested, genotyped for the targeted gene, noted for arrest at a particular embryonic developmental stage, and documented for changes in phenotype. The presence of the *LacZ* cassette in the targeted gene will be useful in obtaining thorough expression profiling during development and/or in specific tissues, by staining for ß-galactosidase activity and/or western blot analysis with a ß-galactosidase antibody.

REFERENCES

Abdel-Meguid N, Ashour AM. Holoprosencephaly and split hand/foot: an additional case with this rare association. *Clin Dysmorphol* 2001; **10**:277-279.

Acebron A, Aza-Blanc P, Rossi DL, *et al.* Congenital human thyroglobulin defect due to low expression of the thyroid-specific transcription factor TTF-1. *J Clin Invest* 1995; **96**:781-785.

Adamson ED, Minchiotti G, Salomon DS. Cripto: a tumor growth factor and more. *J Cell Physiol* 2002; **190**:267-278.

Amendt BA, Sutherland LB, Semina EV, *et al.* The molecular basis of Rieger syndrome. Analysis of Pitx2 homeodomain protein activities. *J Biol Chem* 1998; **273**:20066-20072.

Arakawa H, Nakamura T, Zhadanov AB *et al.* Identification and characterization of the ARP1 gene, a target for the human acute leukemia ALL1 gene. *Proc Natl Acad Sci U S A* 1998; **95**:4573-4578.

Armbruster-Moraes E, Schultz R, Brizot MD, *et al.* Holoprosencephaly in a Klinefelter fetus. *Am J Med Genet* 1999; **85**:511-512.

Aruga J, Yozu A, Hayashizaki Y, *et al.* Identification and characterization of Zic4, a new member of the mouse Zic gene family. *Gene* 1996; **172**:291-294.

Ashworth A, Rastan S, Lovell-Badge R, Kay G X-chromosome inactivation may explain the difference in viability of XO humans and mice. Nature 1991; **351**:406-408.

Back E, Stier R, Bohm N, *et al.* Partial monosomy 22pter leads to q11 in a newborn with the clinical features of trisomy 13 syndrome. *Ann Genet* 1980;**23**:244-248.

Barkovich AJ, Quint DJ. Middle interhemispheric fusion: an unusual variant of holoprosencephaly. *AJNR Am J Neuroradiol* 1993; **14**:431-440.

Barr M Jr, Hanson JW, Currey K, *et al.* Holoprosencephaly in infants of diabetic mothers. *J Pediatr* 1983; **102**:565-568.

Becktor KB, Sverrild L, Pallisgaard C, *et al.* Eruption of the central incisor, the intermaxillary suture, and maxillary growth in patients with a single median maxillary central incisor. *Acta Odontol Scand* 2001; **59**:361-366.

Belloni E, Muenke M, Roessler E, *et al.* Identification of Sonic hedgehog as a candidate gene responsible for holoprosencephaly. *Nat Genet* 1996; **14**:353-356.

Benedyk MJ, Mullen JR, DiNardo S. odd-paired: a zinc finger pair-rule protein required for the timely activation of engrailed and wingless in Drosophila embryos. *Genes Dev* 1994; **8**:105-117.

Bernard JP, Drummond CL, Zaarour P, *et al.* A new clue to the prenatal diagnosis of lobar holoprosencephaly: the abnormal pathway of the anterior cerebral artery crawling under the skull. *Ultrasound Obstet Gynecol* 2002; **19**:605-607.

Bertolino E, Reimund B, Wildt-Perinic D, *et al.* A novel homeobox protein which recognizes a TGT core and functionally interferes with a retinoid-responsive motif. *J Biol Chem* 1995; **270**:31178-31188.

Bingle CD. Thyroid transcription factor-1. *Int J Biochem Cell Biol* 1997; **29**:1471-1473.

Bhat KM. Segment polarity genes in neuroblast formation and identity specification during Drosophila neurogenesis. *Bioessays* 1999; **21**:472-485.

Bouley M, Tian MZ, Paisley K, *et al.* The L1-type cell adhesion molecule neuroglian influences the stability of neural ankyrin in the Drosophila embryo but not its axonal localization. *J Neurosci* 2000; **20**:4515-4523.

Boysen C, Simon MI, Hood L. Analysis of the 1.1-Mb human alpha/delta T-cell receptor locus with bacterial artificial chromosome clones. *Genome Res* 1997; **7**:330-338.

Breedveld GJ, van Dongen JW, Danesino C, *et al.* Mutations in TITF-1 are associated with benign hereditary chorea. *Hum Mol Genet* 2002; **11**:971-979.

Briscoe J, Ericson J. The specification of neuronal identity by graded Sonic Hedgehog signalling. *Semin Cell Dev Biol* 1999; **10**:353-362.

Bronshtein M, Wiener Z. Early transvaginal sonographic diagnosis of alobar holoprosencephaly. *Prenat Diagn* 1991; **11**:459-462.

Brott BK, Sokol SY. Regulation of Wnt/LRP Signaling by Distinct Domains of Dickkopf Proteins. *Mol Cell Biol* 2002; **22**:6100-610.

Brown S, Russo J, Chitayat D, *et al.* The 13q- syndrome: the molecular definition of a critical deletion region in band 13q32. *Am J Hum Genet* 1995; **57**:859-866.

Brown SA, Warburton D, Brown LY, *et al.* Holoprosencephaly due to mutations in ZIC2, a homologue of Drosophila odd-paired. *Nat Genet* 1998; **20**:180-183.

Brown LY, Odent S, David V, *et al.* Holoprosencephaly due to mutations in ZIC2: alanine tract expansion mutations may be caused by parental somatic recombination. *Hum Mol Genet* 2001; **10**:791-796.

Brown LY, Hodge SE, Johnson WG, *et al.* Possible association of NTDs with a polyhistidine tract polymorphism in the ZIC2 gene. *Am J Med Genet* 2002; **108**:128-131.

Brunskill EW, Witte DP, Shreiner AB, *et al.* Characterization of npas3, a novel basic helix-loop-helix PAS gene expressed in the developing mouse nervous system. *Mech Dev* 1999; **88**:237-241.

Bruyere H, Favre B, Douvier S, *et al.* De novo interstitial proximal deletion of 14q and prenatal diagnosis of holoprosencephaly. *Prenat Diagn* 1996; **16**:1059-1060.

Buchanan PD, Rao KW, Doerr CL, *et al.* A complex translocation involving chromosomes 3,11, and 14 with an interstitial deletion, del(14) (q13q22) in a child with congenital glaucoma and cleft lip and palate. *Birth Defects Orig Artic Ser* 1978;**14**:317-322.

Cannistra C, Barbet P, Parisi P, et al. Cyclopia: a radiological and anatomical craniofacial post mortem study. *J Craniomaxillofac Surg* 2001;**29**:150-155.

Carstens MH. Development of the facial midline. *J Craniofac Surg* 2002; **13**:129-190.

Chambon P. The molecular and genetic dissection of the retinoid signalling pathway. *Gene* 1993; **135**:223-228.

Chan W, Kordeli E, Bennett V. 440-kD ankyrinB: structure of the major developmentally regulated domain and selective localization in unmyelinated axons. *J Cell Biol* 1993; **123**:1463-1473.

Chen CP, Lee C-C, Chen L-F, *et al.* Prental diagnosis of de novo proximal interstitial deletion of 14q associated with cebocephaly. *J Med Genet* 1997; **34**:777-778.

Chen CP, Chern SR, Lee CC, *et al.* De novo unbalanced translocation resulting in monosomy for proximal 14q and distal 4p in a fetus with intrauterine growth retardation, Wolf-Hirschhorn syndrome, hypertrophic cardiomyopathy, and partial hemihypoplasia. *J Med Genet* 1998; **35**:1050-1053.

Chen CP, Chern SR, Du SH, *et al.* Molecular diagnosis of a novel heterozygous 268C-->T (R90C) mutation in TGIF gene in a fetus with holoprosencephaly and premaxillary agenesis. *Prenat Diagn* 2002; **22**:5-7.

Chiang C, Litingtung Y, Lee E, *et al.* Cyclopia and defective axial patterning in mice lacking Sonic hedgehog gene function. Nature 1996; **383**:407-413.

Chow BH, Loh SF, Yan YL, *et al.* Holoprosencephaly and chromosomal anomalies. *Singapore Med J* 1996; **37**:394-397.

Chow RL, Altmann CR, Lang RA, *et al.* Pax6 induces ectopic eyes in a vertebrate. *Development* 1999; **126**:4213-4222.

Chrast R, Scott HS, Madani R, *et al.* Mice trisomic for a bacterial artificial chromosome with the single-minded 2 gene (Sim2) show phenotypes similar to some of those present in the partial trisomy 16 mouse models of Down syndrome. *Hum Mol Genet* 2000; **9**:1853-1864.

Cohen MM Jr, Jirasek JE, Guzman RT, *et al.* Holoprosencephaly and facial dysmorphia: nosology, etiology and pathogenesis. *Birth Defects Orig Artic Ser* 1971; **7**:125-135.

Cohen MM Jr. "Holoprosencepahly: cyclopia series" in *Mental Retardation and Congenital Malformations of the Central Nervous system.* Year Book Medical Publishers 1981; p. 1176-1190.

Cohen MM Jr. Perspectives on holoprosencephaly: Part I. Epidemiology, genetics, and syndromology. *Teratology* 1989a; **40**:211-235.

Cohen MM Jr. Perspectives on holoprosencephaly: Part III. Spectra, distinctions, continuities, and discontinuities. *Am J Med Genet* 1989b; **34**:271-288.

Cohen MM Jr. Syndromology: an updated conceptual overview. III. Syndrome delineation. *Int J Oral Maxillofac Surg* 1989c; **18**:281-285.

Cohen MM Jr, Sulik KK. Perspectives on holoprosencephaly: Part II. Central nervous system, craniofacial anatomy, syndrome commentary, diagnostic approach, and experimental studies. *J Craniofac Genet Dev Biol* 1992; **12**:196-244.

Cohen MM Jr. Problems in the definition of holoprosencephaly. *Am J Med Genet* 2001; **103**:183-187.

Cooke A, Tolmie JL, Colgan JM, *et al.* Detection of an unbalanced translocation (4;14) in a mildly retarded father and son by flow cytometry. *Hum Genet* 1989; **83**:83-87.

Corona-Rivera A, Corona-Rivera JR, Bobadilla-Morales L, *et al.* Holoprosencephaly, hypertelorism, and ectrodactyly in a boy with an apparently balanced de novo t(2;4) (q14.2;q35). *Am J Med Genet* 2000; **90**:423-426.

Corsello G, Buttitta P, Cammarata M, *et al.* Holoprosencephaly: examples of clinical variability and etiologic heterogeneity. *Am J Med Genet* 1990; **37**:244-249.

Cox CJ, Espinoza HM, McWilliams B *et al.* Differential regulation of gene expression by PITX2 isoforms. *J Biol Chem* 2002;**277**:25001-25010.

Craddock N, Lendon C. Chromosome Workshop: chromosomes 11, 14, and 15. *Am J Med Genet* 1999; **88**:244-254.

Crews ST. Control of cell lineage-specific development and transcription by bHLH-PAS proteins. *Genes Dev* 1998; **12**:607-620.

Crews ST, Fan CM. Remembrance of things PAS: regulation of development by bHLH-PAS proteins. *Curr Opin Genet Dev* 1999; **9**:580-587.

Croen LA, Shaw GM, Lammer EJ. Risk factors for cytogenetically normal holoprosencephaly in California: a population-based case-control study. *Am J Med Genet* 2000; **90**:320-325.

Das P, Stockton DW, Bauer C *et al.* Haploinsufficiency of PAX9 is associated with autosomal dominant hypodontia. *Hum Genet* 2002 **110**:371-376.

Datson NA, Semina E, van Staalduinen AA *et al.* Closing in on the Rieger syndrome gene on 4q25: mapping translocation breakpoints within a 50-kb region. *Am J Hum Genet* 1996; **59**:1297-1305.

De La Cruz JM, Bamford RN, *et al.* A loss-of-function mutation in the CFC domain of TDGF1 is associated with human forebrain defects. *Hum Genet* 2002; **110**:422-428.

Degar BA, Baskaran N, Hulspas R, *et al.* The homeodomain gene Pitx2 is expressed in primitive hematopoietic stem/progenitor cells but not in their differentiated progeny. *Exp Hematol* 2001; **29**:894-902.

Deloukas P, Schuler GD, Gyapay G, *et al.* A physical map of 30,000 human genes. *Science* 1998; **282**:744-746.

DeMyer W, Zeman W. Alobar holoprosencephaly (arhinencephaly) with median cleft lip and palate. Clinical, nosologic and electroencephalographic considerations. *Confi. Neurol.* 1963; **23**:1-36.

DeMyer W, Zeman W, Palmer CG. The face predicts the brain: Diagnostic significance of the median facial anomalies for holoprosencephaly (arhinencephaly). *Pediatrics* 1964 **34**:257-263

Devriendt K, Vanhole C, Matthijs G, *et al.* Deletion of thyroid transcription factor-1 gene in an infant with neonatal thyroid dysfunction and respiratory failure. *N Engl J Med* 1998a **338**:1317-1318.

Devriendt K, Fryns JP, Chen CP. Holoprosencephaly in deletions of proximal chromosome 14q. *J Med Genet* 1998b; **35**:612.

Ding J, Yang L, Yan YT, *et al.* Cripto is required for correct orientation of the anterior-posterior axis in the mouse embryo. *Nature* 1998; **395**:702-707.

Dobyns WB. Absence makes the search grow longer. *Am J Hum Genet* 1996;**58**:7-16.

Dono R, Scalera L, Pacifico F, *et al.* The murine cripto gene: expression during mesoderm induction and early heart morphogenesis. *Development* 1993; **118**:1157-1168.

Doward W, Perveen R, Lloyd IC, *et al.* A mutation in the RIEG1 gene associated with Peters' anomaly. *J Med Genet* 1999; **36**:152-155.

Dou CL, Li S, Lai E. Dual role of brain factor-1 in regulating growth and patterning of the cerebral hemispheres. *Cereb Cortex* 1999; **9**:543-550.

Dou C, Lee J, Liu B, *et al.* BF-1 interferes with transforming growth factor beta signaling by associating with Smad partners. *Mol Cell Biol* 2000; **20**:6201-6211.

Elias DL, Kawamoto HK Jr, Wilson LF. Holoprosencephaly and midline facial anomalies: redefining classification and management. *Plast Reconstr Surg* 1992; **90**:951-958.

Epstein DJ, Martinu L, Michaud JL, *et al.* Members of the bHLH-PAS family regulate Shh transcription in forebrain regions of the mouse CNS. *Development* 2000; **127**:4701-4709.

Feldman B, Stemple DL. Morpholino phenocopies of sqt, oep, and ntl mutations. *Genesis* 2001; **30**:175-177.

Ferguson-Smith MA. Genetic analysis by chromosome sorting and painting: phylogenetic and diagnostic applications. Eur J Hum Genet 1997; **5**:253-265.

Fiermonte G, Dolce V, Palmieri L *et al.* Identification of the human mitochondrial oxodicarboxylate carrier. Bacterial expression, reconstitution, functional characterization, tissue distribution, and chromosomal location. *J Biol Chem* 2001; **276**:8225-8230.

Fix JD, Dudek RW. "Nervous system" and "Head and neck" in Embryology. Williams and Wilkins. 1995; Chapter 7:75-94, Chapter 12:145-156.

Flomen RH, Vatcheva R, Gorman PA *et al.* Construction and analysis of a sequence-ready map in 4q25: Rieger syndrome can be caused by haploinsufficiency of RIEG, but also by chromosome breaks approximately 90 kb upstream of this gene. *Genomics* 1998; **47**:409-413.

Forrester MB, Merz RD. Epidemiology of holoprosencephaly in Hawaii, 1986-97. *Paediatr Perinat Epidemiol* 2000; **14**:61-63.

Fryns JP, Kleczkowska A, Kenis H. De novo complex chromosomal rearrangement (CCR) in a severely mentally retarded boy. *Ann Genet* 1984;**27**:62-64.

Gage PJ, Suh H, Camper SA. Dosage requirement of Pitx2 for development of multiple organs. *Development* 1999; **126**:4643-4651.

Gebbia M, Ferrero GB, Pilia G, *et al.* X-linked situs abnormalities result from mutations in ZIC3. *Nat Genet* 1997; **17**:305-308.

Geschwind DH, Loginov M, Stern JM. Identification of a locus on chromosome 14q for idiopathic basal ganglia calcification (Fahr disease). *Am J Hum Genet* 1999; **65**:764-772.

Gofflot F, Gaoua W, Bourguignon L, *et al.* Expression of Sonic Hedgehog downstream genes is modified in rat embryos exposed in utero to a distal inhibitor of cholesterol biosynthesis. *Dev Dyn* 2001; **220**:99-111.

Golden JA, Bracilovic A, McFadden KA, *et al.* Ectopic bone morphogenetic proteins 5 and 4 in the chicken forebrain lead to cyclopia and holoprosencephaly. *Proc Natl Acad Sci U S A* 1999; **96**:2439-2444.

Goodrich LV, Milenkovic L, Higgins KM, *et al.* Altered neural cell fates and medulloblastoma in mouse patched mutants. *Science* 1997; **277**:1109-1113.

Goumans MJ, Mummery C. Functional analysis of the TGFbeta receptor/Smad pathway through gene ablation in mice. *Int J Dev Biol* 2000; **44**:253-265.

Govaerts L, Toorman J, Blij-Philipsen MV, *et al.* Another patient with a deletion 14q11.2q13. *Ann Genet* 1996;**39**:197-200.

Grammatico P, de Sanctis S, di Rosa C, *et al.* First case of deletion 14q11.2q13: clinical phenotype. *Ann Genet* 1994;**37**:30-32.

Green PD, Hjalt TA, Kirk DE *et al* Antagonistic regulation of Dlx2 expression by PITX2 and Msx2: implications for tooth development. *Gene Expr* 2001; **9**:265-281.

Gripp KW, Wotton D, Edwards MC, *et al.* Mutations in TGIF cause holoprosencephaly and link NODAL signalling to human neural axis determination. *Nat Genet* 2000; **25**:205-208.

Gritsman K, Zhang J, Cheng S, *et al.* The EGF-CFC protein one-eyed pinhead is essential for nodal signaling. *Cell* 1999; **97**:121-132.

Gurrieri F, Trask BJ, van den Engh G, *et al.* Physical mapping of the holoprosencephaly critical region on chromosome 7q36. *Nat Genet* 1993; **3**:247-251.

Hahn H, Wojnowski L, Miller G, *et al.* The patched signaling pathway in tumorigenesis and development: lessons from animal models. *J Mol Med* 1999; **77**:459-468.

Hammerschmidt M, Brook A, McMahon AP. The world according to hedgehog. *Trends Genet* 1997; **13**:14-21.

Haskill S, Beg AA, Tompkins SM, *et al.* Characterization of an immediate-early gene induced in adherent monocytes that encodes I kappa B-like activity. *Cell* 1991; **65**:1281-1289.

Hogan B, Beddington R, Costantini F, *et al.* Manipulating the Mouse Embryo: A laboratory manual. *Cold Spring Harbor Laboratory Press* 1994: Second edition.

Helms JA, Kim CH, Hu D, *et al.* Sonic hedgehog participates in craniofacial morphogenesis and is down-regulated by teratogenic doses of retinoic acid. *Dev Biol* 1997; **187**:25-35.

Hennekam RC, Van Noort G, de la Fuente FA, *et al.* Agenesis of the nasal septal cartilage: another sign in autosomal dominant holoprosencephaly. *Am J Med Genet* 1991; **39**:121-122.

Hercig D, Bau A, Reznek L. Psychosis associated with lobar holoprosencephaly. Can J Psychiatry 1994;**39**:449.

Hill CS. TGF-beta signalling pathways in early Xenopus development. *Curr Opin Genet Dev* 2001; **11**:533-540.

Hjalt TA, Amendt BA, Murray JC. PITX2 regulates procollagen lysyl hydroxylase (PLOD) gene expression: implications for the pathology of Rieger syndrome. *J Cell Biol* 2001; **152**:545-552.

Hudgins RJ, Edwards MS, Ousterhout DK, *et al.* Pediatric neurosurgical implications of the amniotic band disruption complex. Case reports and review of the literature. *Pediatr Neurosci* 1985-86;**12**:232-239.

Hudson TJ, Stein LD, Gerety SS, *et al.* An STS-based map of the Human genome. *Science* 1995; **270**:1945-1954.

Huh S, Hatini V, Marcus RC, *et al.* Dorsal-ventral patterning defects in the eye of BF-1-deficient mice associated with a restricted loss of shh expression. *Dev Biol* 1999; **211**:53-63.

Hsu TY, Chang SY, Ou CY, *et al.* First trimester diagnosis of holoprosencephaly and cyclopia with triploidy by transvaginal three-dimensional ultrasonography. *Eur J Obstet Gynecol Reprod Biol* 2001; **96**:235-237.

Ikeda K, Shaw-White JR, Wert SE, *et al.* Hepatocyte nuclear factor 3 activates transcription of thyroid transcription factor 1 in respiratory epithelial cells. *Mol Cell Biol* 1996; **16**:3626-3636.

Imaizumi K, Ishii T, Masuno M, *et al.* Association of holoprosencephaly, ectrodactyly, cleft lip/cleft palate and hypertelorism: a possible third case. *Clin Dysmorphol* 1998; **7**:213-216.

Ingham PW, McMahon AP. Hedgehog signaling in animal development: paradigms and principles. *Genes Dev* 2001; **15**:3059-3087.

Ishibashi Y, Maita H, Yano M, *et al.* Pim-1 translocates sorting nexin 6/TRAF4-associated factor 2 from cytoplasm to nucleus. *FEBS Lett* 2001; **506**:33-38.

Iwatani N, Mabe H, Devriendt K, *et al.* Deletion of NKX2.1 gene encoding thyroid transcription factor-1 in two siblings with hypothyroidism and respiratory failure. *J Pediatr* 2000; **137**:272-276.

Jellinger K, Gross H, Kaltenback E, *et al.* Holoprosencephaly and agenesis of the corpus callosum: frequency of associated malformations. *Acta Neuropathol (Berl)* 1981; **55**:1-10.

Jones MH, Hamana N, Nezu J, *et al.* A novel family of bromodomain genes. *Genomics* 2000; **63**:40-45.

Kahn A. Craniofacial anomalies: A Beginner's Guide for Speech-Language Pathologists. Singular publishing group Thomas Learning 2000.

Kamnasaran D, O'Brien PC, Schuffenhauer S, *et al.* Defining the breakpoints of proximal chromosome 14q rearrangements in nine patients using flow-sorted chromosomes. *Am J Med Genet* 2001; **102**:173-182.

Kamnasaran D, Cox DW. Current status of human chromosome 14. *J Med Genet* 2002;**39**:81-90.

Kannan S, De Santis M, Lohmeyer M, *et al.* Cripto enhances the tyrosine phosphorylation of Shc and activates mitogen-activated protein kinase (MAPK) in mammary epithelial cells. *J Biol Chem* 1997; **272**:3330-3335.

Karantanas AH, Papanikolaou N, Danos A, *et al.* Cyclopia and exadactyly: CT and MRI findings. *Dentomaxillofac Radiol* 1999; **28**:372-374.

Kato M, Nanba E, Akaboshi S, *et al.* Sonic hedgehog signal peptide mutation in a patient with holoprosencephaly. *Ann Neurol* 2000; **47**:514-516.

Kawakami K, Ohto H, Takizawa T, *et al.* Identification and expression of six family genes in mouse retina. *FEBS Lett* 1996; **393**:259-263.

Kawakami K, Sato S, Ozaki H, *et al.* Six family genes--structure and function as transcription factors and their roles in development. *Bioessays* 2000; **22**:616-626.

Kim KC, Wakui K, Yamagishi A, *et al* Tetramelic mirror-image polydactyly and a de novo balanced translocation between 2p23.3 and 14q13. *Am J Med Genet* 1997;**68**:70-73.

Kimura S, Hara Y, Pineau T *et al.* The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. *Genes Dev* 1995; **10**:60-69.

Kitamura K, Miura H, Miyagawa-Tomita S *et al.* Mouse Pitx2 deficiency leads to anomalies of the ventral body wall, heart, extra- and periocular mesoderm and right pulmonary isomerism. *Development* 1999; **126**:5749-5758.

Kjaer I, Keeling JW, Fischer Hansen B, *et al.* Midline skeletodental morphology in holoprosencephaly. *Cleft Palate Craniofac J* 2002; **39**:357-363.

Kleinjan DJ, van Heyningen V. Position effect in human genetic disease. *Hum Mol Genet* 1998;**7**:1611-8.

Kobayashi M, Toyama R, Takeda H, *et al.* Overexpression of the forebrain-specific homeobox gene six3 induces rostral forebrain enlargement in zebrafish. *Development* 1998; **125**:2973-2982.

Kondoh S, Sugawara H, Harada N, *et al.* A novel gene is disrupted at a 14q13 breakpoint of t(2;14) in a patient with mirror-image polydactyly of hands and feet. *J Hum Genet* 2002;**47**:136-139.

Koyabu Y, Nakata K, Mizugishi K, *et al.* Physical and functional interactions between Zic and Gli proteins. *J Biol Chem* 2001; **276**:6889-6892.

Kozak M. Interpreting cDNA sequences: some insights from studies on translation. *Mamm Genome* 1996; **7**:563-574.

Kozlowski K, Walter MA. Variation in residual PITX2 activity underlies the phenotypic spectrum of anterior segment developmental disorders. *Hum Mol Genet* 2000; **9**:2131-2139.

Kraynack NC, Hostoffer RW, Robin NH. Agenesis of the corpus callosum associated with DiGeorge-velocardiofacial syndrome: a case report and review of the literature. *J Child Neurol* 1999; **14**:754-756.

Krude H, Schutz B, Biebermann H, *et al.* Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency. *J Clin Invest* 2002; **109**:475-480.

Kulak SC, Kozlowski K, Semina EV, *et al.* Mutation in the RIEG1 gene in patients with iridogoniodysgenesis syndrome. *Hum Mol Genet* 1998; **7**:1113-1117.

Kulharya AS, Maberry M, Kukolich MK *et al.* Interstitial deletions 4q21.1q25 and 4q25q27: phenotypic variability and relation to Rieger anomaly. *Am J Med Genet* 1995; **55**:165-170.

Lahbib-Mansais Y, Dalias G, Milan D, *et al.* A successful strategy for comparative mapping with human ESTs: 65 new regional assignments in the pig. *Mamm Genome* 1999; **10**:145-153.

Lapunzina P, Musante G, Pedraza A, *et al.* Semilobar holoprosencephaly, coronal craniosynostosis, and multiple congenital anomalies: a severe expression of the Genoa syndrome or a newly recognized syndrome? *Am J Med Genet* 2001; **102**:258-260.

Larsen WJ. "Development of the head, the neck and the eyes and ears" and "Development of the brain and cranial nerves" in Human embryology. Churchill Livingstone. 1993; Chapter 12:309-367, Chapter 13:375-410.

Leppert GS, Yang JM, Sundin OH. Sequence and location of SIX3, a homeobox gene expressed in the human eye. *Ophthalmic Genet* 1999; **20**:7-21.

Levin SW, Surana RB. 1991. Holoprosencephaly associated with 46,XX,del(14)(q11.1q13) Am J Hum Genet Suppl 1991;**49**:269.

Li H, Tao W, Lai E. Characterization of the structure and function of the gene for transcription factor BF-1, an essential regulator of forebrain development. *Brain Res Mol Brain Res* 1996; **37**:96-104.

Lin CR, Kioussi C, O'Connell S *et al.* Pitx2 regulates lung asymmetry, cardiac positioning and pituitary and tooth morphogenesis. *Nature* 1999; **401**:279-282.

Livy DJ, Schalomon PM, Roy M, *et al.* Increased axon number in the anterior commissure of mice lacking a corpus callosum. *Exp Neurol* 1997; **146**:491-501.

Lo RS, Wotton D, Massague J. Epidermal growth factor signaling via Ras controls the Smad transcriptional co-repressor TGIF. *EMBO J* 2001; **20**:128-136.

Lu MF, Pressman C, Dyer R, *et al.* Function of Rieger syndrome gene in left right asymmetry and craniofacial development. *Nature* 1999; **401**:276-278.

Luo K, Stroschein SL, Wang W, *et al.* The Ski oncoprotein interacts with the Smad proteins to repress TGFbeta signaling. *Genes Dev* 1999; **13**:2196-2206.

Makita Y, Masuno M, Imaizumi K *et al.* Rieger syndrome with de novo reciprocal translocation t(1;4) (q23.1;q25). *Am J Med Genet* 1995; **57**:19-21.

Mandelli F, Avvisati G, Lo Coco F. Advances in the understanding and management of acute promyelocytic leukemia. *Rev Clin Exp Hematol* 2002; **6**:60-71, 86-87.

Mao B, Wu W, Davidson G, *et al.* Kremen proteins are Dickkopf receptors that regulate Wnt/beta-catenin signalling. *Nature* 2002; **417**:664-667.

Mao B, Wu W, Li Y, *et al.* LDL-receptor-related protein 6 is a receptor for Dickkopf proteins. *Nature* 2001; **411**:321-325.

Marcorelles P, Loget P, Fallet-Bianco C, *et al.* Unusual variant of holoprosencephaly in monosomy 13q. *Pediatr Dev Pathol* 2002; **5**:170-178.

Marek A, Brodzicki J, Liberek A, *et al.* TGF-b (transforming growth factor -b) in chronic inflammatory conditions - a new diagnostic and prognostic marker? *Med Sci Monit* 2002; **8**:RA145-51.

Marigo V, Roberts DJ, Lee SM, *et al.* Cloning, expression, and chromosomal location of SHH and IHH: two human homologues of the Drosophila segment polarity gene hedgehog. *Genomics* 1995;**28**:44-51.

Mark M, Ghyselinck NB, Kastner P, *et al.* Mesectoderm is a major target of retinoic acid action. *Eur J Oral Sci* 1998; 106:24-31.

Matsumoto N, Soeda E, Ohashi H *et al* A 1.2 megabase BAC/PAC contig spanning the 14q13 breakpoint of t(2;14) in a mirror-image polydactyly patient. *Genomics* 1997;**45**:11-16.

Matsunaga E, Shiota K. Holoprosencephaly in human embryos: epidemiologic studies of 150 cases. *Teratology* 1977; **16**:261-272.

Mehta L, Cervantes C, Small K, *et al.* Proximal deletion of chromosome 14q. *Am J Hum Genet Suppl* 1999;**65**:351.

Melhuish TA, Gallo CM, Wotton D. TGIF2 interacts with histone deacetylase 1 and represses transcription. *J Biol Chem* 2001; **276**:32109-32114.

Minchiotti G, Parisi S, Liguori GL, *et al.* Role of the EGF-CFC gene cripto in cell differentiation and embryo development. *Gene* 2002; **287**:33-37.

Ming JE, Muenke M. Holoprosencephaly: from Homer to Hedgehog. *Clin Genet* 1998; **53**:155-163.

Ming JE, Kaupas ME, Roessler E, *et al.* Mutations in PATCHED-1, the receptor for SONIC HEDGEHOG, are associated with holoprosencephaly. *Hum Genet* 2002; **110**:297-301.

Ming JE, Muenke M. Multiple hits during early embryonic development: digenic diseases and holoprosencephaly. *Am J Hum Genet* 2002; **71**:1017-1031.

Miyake T, Cameron AM, Hall BK. Detailed staging of inbred C57BL/6 mice between Theiler's [1972] stages 18 and 21 (11-13 days of gestation) based on craniofacial development. *J Craniofac Genet Dev Biol* 1996; **16**:1-31.

Mizugishi K, Aruga J, Nakata K, *et al.* Molecular properties of Zic proteins as transcriptional regulators and their relationship to GLI proteins. *J Biol Chem* 2001; **276**:2180-2188.

Mohn AR, Gainetdinov RR, Caron MG, *et al.* Mice with reduced NMDA receptor expression display behaviors related to schizophrenia. *Cell* 1999;**98**:427-436.

Moog U, De Die-Smulders CE, Schrander-Stumpel CT, *et al.* Holoprosencephaly: the Maastricht experience. *Genet Couns* 2001;**12**:287-298.

Moore. Preparation of genomic DNA from mammalian tissue. *Current Protocols in Molecular Biology* 1999 **1**:2.2.1-2.2.2.

Muccielli ML, Martinez S, Pattyn A, *et al.* Otlx2, an Otx-related homeobox gene expressed in the pituitary gland and in a restricted pattern in the forebrain. *Mol Cell Neurosci* 1996; **8**:258-271.

Mucchielli ML, Mitsiadis TA, Raffo S, *et al.* Mouse Otlx2/RIEG expression in the odontogenic epithelium precedes tooth initiation and requires mesenchyme-derived signals for its maintenance. *Dev Biol* 1997; **189**:275-284.

Muenke M, Gurrieri F, Bay C, *et al.* Linkage of a human brain malformation, familial holoprosencephaly, to chromosome 7 and evidence for genetic heterogeneity. *Proc Natl Acad Sci U S A* 1994; **91**:8102-8106.

Muenke M, Bone LJ, Mitchell HF, *et al.* Physical mapping of the holoprosencephaly critical region in 21q22.3, exclusion of SIM2 as a candidate gene for holoprosencephaly, and mapping of SIM2 to a region of chromosome 21 important for Down syndrome. *Am J Hum Genet* 1995; **57**:1074-1079.

Muenke M, Beachy PA. Genetics of ventral forebrain development and holoprosencephaly. *Curr Opin Genet Dev* 2000; **10**:262-269.

Mukhopadhyay M, Shtrom S, Rodriguez-Esteban C, *et al.* Dickkopf1 is required for embryonic head induction and limb morphogenesis in the mouse. *Dev Cell* 2001; **1**:423-434.

Muller F, O'Rahilly R. Mediobasal prosencephalic defects, including holoprosencephaly and cyclopia, in relation to the development of the human forebrain. *Am J Anat* 1989; **185**:391-414.

Muthalif MM, Karzoun NA, Benter IF, *et al.* Functional significance of activation of calcium/calmodulin-dependent protein kinase II in angiotensin II--induced vascular hyperplasia and hypertension. *Hypertension* 2002; **39**:704-709.

Nagai T, Aruga J, Takada S, *et al.* The expression of the mouse Zic1, Zic2, and Zic3 gene suggests an essential role for Zic genes in body pattern formation. *Dev Biol* 1997; **182**:299-313.

Nagai T, Aruga J, Minowa O, *et al.* Zic2 regulates the kinetics of neurulation. *Proc Natl Acad Sci U S A* 2000; **97**:1618-1623.

Nakata K, Koyabu Y, Aruga J, *et al.* A novel member of the Xenopus Zic family, Zic5, mediates neural crest development. *Mech Dev* 2000; **99**:83-91.

Nambu JR, Lewis JO, Wharton KA Jr, *et al.* The Drosophila single-minded gene encodes a helix-loop-helix protein that acts as a master regulator of CNS midline development. *Cell* 1991; **67**:1157-1167.

Nanni L, Ming JE, Bocian M, *et al.* The mutational spectrum of the sonic hedgehog gene in holoprosencephaly: SHH mutations cause a significant proportion of autosomal dominant holoprosencephaly. *Hum Mol Genet* 1999; **8**:2479-2488.

Nanni L, Croen LA, Lammer EJ, *et al.* Holoprosencephaly: molecular study of a California population. *Am J Med Genet* 2000; **90**:315-319.

Nanni L, Ming JE, Du Y, *et al.* SHH mutation is associated with solitary median maxillary central incisor: a study of 13 patients and review of the literature. *Am J Med Genet* 2001; **102**:1-10.

Niehrs C, Kazanskaya O, Wu W, *et al.* Dickkopf1 and the Spemann-Mangold head organizer. *Int J Dev Biol* 2001;**45**:237-40.

Nusslein-Volhard C, Wieschaus E. Mutations affecting segment number and polarity in Drosophila. *Nature* 1980;**287**:795-801.

Odent S, Le Marec B, Munnich A, *et al*. Segregation analysis in nonsyndromic holoprosencephaly. *Am J Med Genet* 1998; **77**:139-143.

Odent S, Atti-Bitach T, Blayau M, *et al.* Expression of the Sonic hedgehog (SHH) gene during early human development and phenotypic expression of new mutations causing holoprosencephaly. *Hum Mol Genet* 1999; **8**:1683-1689.

Ohyashiki K, Ohyashiki JH, Hojo H, et al. Cytogenetic findings in adult acute leukemia and myeloproliferative disorders with an involvement of megakaryocyte lineage. *Cancer* 1990;**65**:940-948.

Ogilvie CM, Raymond FL, Harrison RH, *et al.* A new approach to the elucidation of complex chromosome rearrangements illustrated by a case of Rieger syndrome. *J Med Genet* 1998; **35**:234-237.

Ogura H, Aruga J, Mikoshiba K. Behavioral abnormalities of Zic1 and Zic2 mutant mice: implications as models for human neurological disorders. *Behav Genet* 2001; **31**:317-324.

Oliver G, Mailhos A, Wehr R, *et al.* Six3, a murine homologue of the sine oculis gene, demarcates the most anterior border of the developing neural plate and is expressed during eye development. *Development* 1995;**121**:4045-4055.

Oliver G, Loosli F, Koster R, *et al.* Ectopic lens induction in fish in response to the murine homeobox gene Six3. *Mech Dev* 1996; **60**:233-239.

Olsen CL, Hughes JP, Youngblood LG, *et al.* Epidemiology of holoprosencephaly and phenotypic characteristics of affected children: New York State, 1984-1989. *Am J Med Genet* 1997; **73**:217-226.

Onda M, Inui S, Maeda K, *et al.* Expression and chromosomal localization of the human alpha 4/IGBP1 gene, the structure of which is closely related to the yeast TAP42 protein of the rapamycin-sensitive signal transduction pathway. *Genomics* 1997; **46**:373-378.

Orioli IM, Castilla EE, Ming JE, *et al.* Identification of novel mutations in SHH and ZIC2 in a South American (ECLAMC) population with holoprosencephaly. *Hum Genet* 2001; **109**:1-6.

Orioli IM, Vieira AR, Castilla EE, *et al.* Mutational analysis of the Sonic Hedgehog gene in 220 newborns with oral clefts in a South American (ECLAMC) population. *Am J Med Genet* 2002; **108**:12-15.

Overdier DG, Porcella A, Costa RH. The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino acid residues adjacent to the recognition helix. *Mol Cell Biol* 1994; **14**:2755-2766.

Overhauser J, Mitchell HF, Zackai EH, *et al.* Physical mapping of the holoprosencephaly critical region in 18p11.3. *Am J Hum Genet* 1995; **57**:1080-1085.

Pabst O, Herbrand H, Arnold HH. Nkx2-9 is a novel homeobox transcription factor which demarcates ventral domains in the developing mouse CNS. *Mech Dev* 1998; **73**:85-93.

Park JP, Moeschler JB, Berg SZ, *et al.* Schizophrenia and mental retardation in an adult male with a de novo interstitial deletion 9(q32q34.1). *J Med Genet* 1991; **28**:282-283.

Parks WT, Frank DB, Huff C, *et al.* Sorting nexin 6, a novel SNX, interacts with the transforming growth factor-beta family of receptor serine-threonine kinases. *J Biol Chem* 2001; **276**:19332-19339.

Pasquier L, Dubourg C, Blayau M, *et al.* A new mutation in the six-domain of SIX3 gene causes holoprosencephaly. *Eur J Hum Genet* 2000; **8**:797-800.

Pauken CM, LaBorde JB, Bolon B. Retinoic acid acts during peri-implantational development to alter axial and brain formation. *Anat Embryol (Berl)* 1999; **200**:645-655.

Peebles DM. Holoprosencephaly. *Prenat Diagn* 1998; **18**:477-480.

Pepinsky RB, Zeng C, Wen D, *et al.* Identification of a palmitic acid-modified form of human Sonic hedgehog. *J Biol Chem* 1998; **273**:14037-14045.

Perler FB. Protein splicing of inteins and hedgehog autoproteolysis: structure, function, and evolution. *Cell* 1998; **92**:1-4.

Perna MG, Civitareale D, de Filippis V, *et al.* Absence of mutations in the gene encoding thyroid transcription factor-1 (TTF-1) in patients with thyroid dysgenesis. *Thyroid* 1997; **7**:377-381.

Pessah M, Prunier C, Marais J, *et al.* c-Jun interacts with the corepressor TG-interacting factor (TGIF) to suppress Smad2 transcriptional activity. *Proc Natl Acad Sci U S A* 2001; **98**:6198-6203.

Platt KA, Michaud J, Joyner AL. Expression of the mouse Gli and Ptc genes is adjacent to embryonic sources of hedgehog signals suggesting a conservation of pathways between flies and mice. *Mech Dev* 1997; **62**:121-135.

Priston M, Kozlowski K, Gill D *et al.* Functional analyses of two newly identified PITX2 mutants reveal a novel molecular mechanism for Axenfeld-Rieger syndrome. *Hum Mol Genet* 2001; **10**:1631-1638.

Ramelli GP, Remonda L, Lovblad KO, *et al.* Abnormal myelination in a patient with deletion 14q11.2q13.1. *Pediatr Neurol* 2000; **23**:170-172.

Rana MW. "Nervous system" and "Craniofacial development" in Embryology made easy. Harwood academic publishers. 1998; Chapter 12: 179-190, Chapter 18: 313-347.

Rasmussen SA, Moore CA, Khoury MJ, *et al.* Descriptive epidemiology of holoprosencephaly and arhinencephaly in metropolitan Atlanta, 1968-1992. *Am J Med Genet* 1996; **66**:320-333.

Rebagliati MR, Toyama R, Haffter P, *et al.* cyclops encodes a nodal-related factor involved in midline signaling. *Proc Natl Acad Sci U S A* 1998; **95**:9932-9937.

Rhinn M, Dierich A, Shawlot W, *et al.* Sequential roles for Otx2 in visceral endoderm and neuroectoderm for forebrain and midbrain induction and specification. *Development* 1998; **125**:845-856.

Riley BP, Tahir E, Rajagopalan S, *et al.* A linkage study of the N-methyl-D-aspartate receptor subunit gene loci and schizophrenia in southern African Bantu-speaking families. *Psychiatr Genet* 1997;**7**:57-74.

Roach E, Demyer W, Conneally PM, *et al.* Holoprosencephaly: birth data, genetic and demographic analyses of 30 families. *Birth Defects Orig Artic Ser* 1975;**11**:294-313.

Robin NH, Ko LM, Heger S *et al.* Syntelencephaly in an infant of a diabetic mother. Am J Med Genet. 1996;**66**:433-437.

Rodriguez C, Huang LJ, Son JK, *et al.* Functional cloning of the proto-oncogene brain factor-1 (BF-1) as a Smad-binding antagonist of transforming growth factor-beta signaling. *J Biol Chem* 2001; **276**:30224-30230.

Roelink H, Porter JA, Chiang C, *et al.* Floor plate and motor neuron induction by different concentrations of the amino-terminal cleavage product of sonic hedgehog autoproteolysis. Cell 1995; **81**:445-455.

Roessler E, Belloni E, Gaudenz K, *et al.* Mutations in the human Sonic Hedgehog gene cause holoprosencephaly. *Nat Genet* 1996; **14**:357-360.

Roessler E, Belloni E, Gaudenz K, *et al.* Mutations in the C-terminal domain of Sonic Hedgehog cause holoprosencephaly. *Hum Mol Genet* 1997; **6**:1847-1853.

Roessler E, Muenke M. Holoprosencephaly: a paradigm for the complex genetics of brain development. *J Inherit Metab Dis* 1998; **21**:481-497.

Roessler E, Mittaz L, Du Y, *et al.* Structure of the human Lanosterol synthase gene and its analysis as a candidate for holoprosencephaly (HPE1). *Hum Genet* 1999; **105**:489-495.

Roessler E, Du Y, Glinka A, *et al.* The genomic structure, chromosome location, and analysis of the human DKK1 head inducer gene as a candidate for holoprosencephaly. *Cytogenet Cell Genet* 2000; **89**:220-224.

Roessler E, Muenke M. Midline and laterality defects: left and right meet in the middle. *Bioessays* 2001; **23**:888-900.

Rohr KB, Schulte-Merker S, Tautz D. Zebrafish zic1 expression in brain and somites is affected by BMP and hedgehog signalling. *Mech Dev* 1999; **85**:147-159.

Rollins N, Joglar J, Perlman J. Coexistent holoprosencephaly and Chiari II malformation. *AJNR Am J Neuroradiol* 1999; **20**:1678-1681.

Rooke HM, Crosier KE. The smad proteins and TGFbeta signalling: uncovering a pathway critical in cancer. *Pathology* 2001; **33**:73-84.

Rubinstein D, Cajade-Law, Youngman V, *et al.* The development of the corpus callosum in semilobar and lobar holoprosencephaly. *Pediatr Radiol* 1996;**26**:839-844.

Ruwhof C, van der Laarse A. Mechanical stress-induced cardiac hypertrophy: mechanisms and signal transduction pathways. *Cardiovasc Res* 2000; **47**:23-37.

Salen G, Shefer S, Batta AK, *et al.* Abnormal cholesterol biosynthesis in the Smith-Lemli-Opitz syndrome. *J Lipid Res* 1996; **37**:1169-1180.

Sarnat HB, Flores-Sarnat L. A new classification of malformations of the nervous system: an integration of morphological and molecular genetic criteria as patterns of genetic expression. *Europ J Paediatr Neurol* 2001;**5**:57-64.

Salero E, Perez-Sen R, Aruga J, *et al.* Transcription factors Zic1 and Zic2 bind and transactivate the apolipoprotein E gene promoter. *J Biol Chem* 2001; **276**:1881-1888.

Schaefer GB, Shuman RM, Wilson DA, *et al.* Partial agenesis of the anterior corpus callosum: correlation between appearance, imaging, and neuropathology. *Pediatr Neurol* 1991; **7**:39-44.

Schell U, Wienberg J, Kohler A, *et al.* Molecular characterization of breakpoints in patients with holoprosencephaly and definition of the HPE2 critical region 2p21. *Hum Mol Genet* 1996; **5**:223-239.

Schier AF, Shen MM. Nodal signalling in vertebrate development. *Nature* 2000; **403**:385-389.

Schinzel A. Incomplete trisomy 22. II. Familial trisomy of the distal segment of chromosome 22q in two brothers from a mother with a translocation, t(6;22)(q27;q13). *Hum Genet* 1981; **56**:263-268.

Schinzel A, Brecevic L, Dutly F, *et al.* Multiple congenital anomalies including the Rieger eye malformation in a boy with interstitial deletion of (4) (q25-->q27) secondary to a balanced insertion in his normal father: evidence for haplotype insufficiency causing the Rieger malformation. *J Med Genet* 1997; **34**:1012-1014.

Schuffenhauer S, Leifheit HJ, Lichtner P, *et al.* De novo deletion (14)(q11.2q13) including PAX9: clinical and molecular findings. *J Med Genet* 1999; **36**:233-236.

Schuler GD, Boguski MS, Hudson TJ, *et al.* The human transcript map. *Science* 1996; **274**:547-562.

Schwarz DG, Griffin CT, Schneider EA, *et al.* Genetic analysis of sorting nexins 1 and 2 reveals a redundant and essential function in mice. *Mol Biol Cell* 2002;**13**(10):3588-600.

Schweickert A, Campionel M, Steinbeisser H, *et al.* Pitx2 isoforms: involvement of Pitx2c but not Pitx2a or Pitx2b in vertebrate left-right asymmetry. *Mech Dev* 2000; **90**:41-51.

Semina EV, Reiter R, Leysens NJ *et al.* Cloning and characterization of a novel bicoid-related homeobox transcription factor gene, RIEG, involved in Rieger syndrome. *Nat Genet* 1996; **14**:392-399.

Semsarian C, Wu MJ, Ju YK, *et al.* Skeletal muscle hypertrophy is mediated by a Ca2+-dependent calcineurin signalling pathway. *Nature* 1999; **400**:576-581.

Sepulveda W, Monckeberg MJ, Be C. Twin pregnancy discordant for trisomy 14 mosaicism: prenatal sonographic findings. *Prenat Diagn* 1998; **18**:481-484.

Shaffer LG, Lupski JR. Molecular mechanisms for constitutional chromosomal rearrangements in humans. *Annu Rev Genet* 2000;**34**:297-329.

Shapira SK, Anderson KL, Orr-Urtregar A, *et al.* De novo proximal interstitial deletions of 14q: cytogenetic and molecular investigations. *Am J Med Genet* 1994; **52**:44-50.

Shiota K, Nakatsu T, Irie H. Computerized three-dimensional reconstruction of the brain of normal and holoprosencephalic human embryos. *Birth Defects Orig Artic Ser* 1993;**29**:261-271.

Siebert JR. The Ethmoid bone: Implications for Normal and Abnormal Facial Development. J Cranio Genet Dev Biol 1981;**1**:381-389.

Signer E, Kuenzle CC, Thomann PE, *et al.* DNA fingerprinting: improved DNA extraction from small blood samples. *Nucleic Acids Res* 1988; **16**:7738.

Simon EM, Hevner RF, Pinter JD, *et al.* The middle interhemispheric variant of holoprosencephaly. *AJNR Am J Neuroradiol* 2002; **23**:151-156.

Smalley MJ, Dale TC. Wnt signalling in mammalian development and cancer. *Cancer Metastasis Rev* 1999;**18**:215-230.

Smilari P, Serra A, La Mantia I, *et al.* Holoprosencephaly (lobar form) associated with bilateral vocal cord palsy. *J Child Neurol* 2001; **16**:932-934.

Soderlund C, Longden I, Mott R. FPC: a system for building contigs from restriction fingerprinted clones. *CABIOS* 1997;**13**:523-535.

Stagiannis KD, Sepulveda W, Bower S. Early prenatal diagnosis of holoprosencephaly: the value of transvaginal ultrasonography. *Eur J Obstet Gynecol Reprod Biol* 1995; **61**:175-176.

Stapleton P, Weith A, Urbanek P, *et al.* Chromosomal localization of seven PAX genes and cloning of a novel family member, PAX-9. *Nat Genet* 1993; **3**:292-298.

Stockton DW, Das P, Goldenberg M, *et al.* Mutation of PAX9 is associated with oligodontia. *Nat Genet* 2000;**24**:18-19.

Stringham HM, Boehnke M, Lange K Point and interval estimates of marker localization in radiation hybrid mapping. *Am J Hum Genet* 1999; **65**:545-553.

Sulik KK, Johnston MC. Embryonic origin of holoprosencephaly: interrelationship of the developing brain and face. *Scan Electron Microsc* 1982;(**Pt 1**):309-322.

Sun SY, Lotan R. Retinoids and their receptors in cancer development and chemoprevention. *Crit Rev Oncol Hematol* 2002; **41**:41-55.

Suthers G, Smith S, Springbett S. Skewed sex ratios in familial holoprosencephaly and in people with isolated single maxillary central incisor. *J Med Genet* 1999; **36**:924-926.

Taipale J, Beachy PA. The Hedgehog and Wnt signalling pathways in cancer. *Nature* 2001; **411**:349-254.

Tavin E, Stecker E, Marion R. Nasal pyriform aperture stenosis and the holoprosencephaly spectrum. *Int J Pediatr Otorhinolaryngol* 1994; **28**:199-204.

Taylor MS. Characterization and comparative analysis of the EGLN gene family. *Gene* 2001; **275**:125-132.

Teasdale RD, Loci D, Houghton F, *et al.* A large family of endosome-localized proteins related to sorting nexin 1. *Biochem J* 2001; **358**:7-16.

Telenius H, Pelmear A, Tunnacliffe A, *et al.* Cytogenetic analysis by chromosome painting using DOP-PCR amplified flow-sorted chromosomes. *Genes Chromosomes Cancer* 1992;**4**:257-263.

Thirion C, Stucka R, Mendel B, *et al.* Characterization of human muscle type cofilin (CFL2) in normal and regenerating muscle. *Eur J Biochem* 2001; **268**:3473-3482.

Thomas JB, Crews ST, Goodman CS. Molecular genetics of the single-minded locus: a gene involved in the development of the Drosophila nervous system. *Cell* 1988; **52**:133-141.

Turner CD, Silva S, Jeanty P. Prenatal diagnosis of alobar holoprosencephaly at 10 weeks of gestation. *Ultrasound Obstet Gynecol* 1999;**13**:360-362.

Urioste M, Valcarcel E, Gomez MA, *et al.* Holoprosencephaly and trisomy 21 in a child born to a nondiabetic mother. *Am J Med Genet* 1988; **30**:925-928.

Vantrappen G, Feenstra L, Fryns JP. On the association profound nerve deafness, semilobar holoprosencephaly, and minor midline developmental anomalies. *Genet Couns* 1999;**10**:399-401.

Vervoort VS, Viljoen D, Smart R, *et al.* Sorting nexin 3 (*SNX3*) is disrupted in a patient with a translocation t(6;13)(q21;q12) and microcephaly, microphthalmia, ectrodactyly, prognathism (MMEP) phenotype. *J Med Genet* 2002; **39**: 893-899.

Villavicencio EH, Walterhouse DO, Iannaccone PM. The sonic hedgehog-patched-gli pathway in human development and disease. *Am J Hum Genet* 2000; **67**:1047-1054.

Voiculescu I, Back E, Duncan AM, *et al.* Trisomy 22 in a newborn with multiple malformations. *Hum Genet* 1987; **76**:298-301.

Wallis DE, Muenke M. Molecular mechanisms of holoprosencephaly. *Mol Genet Metab* 1999; **68**:126-138.

Wallis D, Muenke M. Mutations in holoprosencephaly. *Hum Mutat* 2000;**16**:99-108.

Wallis DE, Roessler E, Hehr U, *et al.* Mutations in the homeodomain of the human SIX3 gene cause holoprosencephaly. *Nat Genet* 1999; **22**:196-198.

Wang Q, Timur AA, Szafranski P, *et al* Identification and molecular characterization of de novo translocation t(8;14)(q22.3;q13) associated with a vascular and tissue overgrowth syndrome. *Cytogenet Cell Genet* 2001a;**95**:183-188.

Wang JH, Kelly P. Calcium-calmodulin signalling pathway up-regulates glutamatergic synaptic function in non-pyramidal, fast spiking rat hippocampal CA1 neurons. *J Physiol* 2001; **533**:407-422.

Warkany J. Syndromes. *Am J Dis Child* 1971; **121**:365-370.

Webster WS, Lipson AH, Sulik KK. Interference with gastrulation during the third week of pregnancy as a cause of some facial abnormalities and CNS defects. *Am J Med Genet* 1988; **31**:505-512.

Wei Q, Adelstein RS. Pitx2a expression alters actin-myosin cytoskeleton and migration of HeLa cells through Rho GTPase signaling. *Mol Biol Cell* 2002; **13**:683-697.

Weinstein M, Yang X, Deng C. Functions of mammalian Smad genes as revealed by targeted gene disruption in mice. *Cytokine Growth Factor Rev* 2000; **11**:49-58.

Whiteford ML, Tolmie JL. Holoprosencephaly in the west of Scotland 1975-1994. *J Med Genet* 1996; **33**:578-584.

Wong HS, Tang MH, Yan KW, *et al.* Histological findings in a case of alobar holoprosencephaly diagnosed at 10 weeks of pregnancy. *Prenat Diagn* 1999; **19**:859-862.

Worby CA and Dixon JE. Sorting out the cellular functions of sorting nexins. *Nat Rev Mol Cell Biol* 2002; **3**:919-31.

Wotton D, Knoepfler PS, Laherty CD, *et al.* The Smad transcriptional corepressor TGIF recruits mSin3. *Cell Growth Differ* 2001; **12**:457-463.

Wotton D, Massague J. Smad transcriptional corepressors in TGF beta family signaling. *Curr Top Microbiol Immunol* 2001;**254**:145-164.

Wraith JE, Super M, Watson GH, *et al.* Velo-cardio-facial syndrome presenting as holoprosencephaly. *Clin Genet* 1985; **27**:408-410.

Xuan S, Baptista CA, Balas G, *et al.* Winged helix transcription factor BF-1 is essential for the development of the cerebral hemispheres. *Neuron* 1995; **14**:1141-1152.

Yan YT, Liu JJ, Luo Y, *et al.* Dual roles of Cripto as a ligand and coreceptor in the nodal signaling pathway. *Mol Cell Biol* 2002; **22**:4439-4449.

Yang YP, Womack JE Parallel radiation hybrid mapping: a powerful tool for high-resolution genomic comparison. *Genome Res* 1998; **8**:731-736.

Yao J, Lai E, Stifani S. The winged-helix protein brain factor 1 interacts with groucho and hes proteins to repress transcription. *Mol Cell Biol* 2001; **21**:1962-1972.

Yasui K, Imoto I, Fukuda Y, *et al.* Identification of target genes within an amplicon at 14q12-q13 in esophageal squamous cell carcinoma. *Genes Chromosomes Cancer* 2001; **32**:112-118.

Yeo C, Whitman M. Nodal signals to Smads through Cripto-dependent and Cripto-independent mechanisms. *Mol Cell* 2001; **7**:949-957.

Yin J, Yang CH, Zwieb C. Assembly of the human signal recognition particle (SRP): overlap of regions required for binding of protein SRP54 and assembly control. *RNA* 2001; **7**:1389-1396.

Young M, Chen H, Lalioti MD, *et al.* The human lanosterol synthase gene maps to chromosome 21q22.3. *Hum Genet* 1996; **97**:620-624.

Yu X, St Amand TR, Wang S *et al.* Differential expression and functional analysis of Pitx2 isoforms in regulation of heart looping in the chick. *Development* 2001; **128**:1005-1013.

APPENDICES

Appendix 1

Contact information for patient specimens (Chapter Two)

Case	Cell line/DNA name	Type	Contact	Reference
1	L251DC	Lymphoblast	Dr. James Lupski (jlupski@bcm.tmc.edu)	Shapira et al., 1994
2	L777DC	Lymphoblast	Dr. James Lupski (jlupski@bcm.tmc.edu)	Shapira et al., 1994
3	L640DC	Lymphoblast	Dr. Paola Grammatico (paola.grammatico@uniroma1.it)	Grammatico et al., 1994
4	HQ0004	Lymphoblast	Dr. Quarrell, European (o.quarrell@shieffeld.ac.uk), European Collection of Cell Cultures	
5	SAN	Lymphoblast	Dr. Simone Schuffenhaeur (simone@pedgen.med.uni-muenchen.de)	Schuffenhaeur et al., 1999
6	MAA	Lymphoblast	Dr. Simone Schuffenhauer (simone@pedgen.med.uni-muenchen.de)	
7	L6873	Lymphoblast	Dr. John Tolmie (j-tolmie.genetics.gla.ac.uk)	Cooke et al., 1989
8	B01060	Lymphoblast	European Collection of Cell Cultures	
9	BV0416	Lymphoblast	European Collection of Cell Cultures	
10	L6504	Lymphoblast	Dr. Maximilan Muenke (mmuenke@nhgri.nih.gov), Hospital for Sick Children, Toronto	
11	SMOM	Lymphoblast	Dr. Walter Muir (wjmuir@srv1.med.ed.ac.uk)	
12	L6874 or SURAM	Lymphoblast	Dr. Walter Muir (wjmuir@srv1.med.ed.ac.uk)	
13	D2502	DNA	Dr. Koenraad Devriendt (Koen.Devriendt@med.kuleuven.ac.be)	Devriendt et al., 1998a
14	LMETHADC	Lymphoblast	Dr. Laskhmi Mehta (davemehta@pool.net), Hospital for Sick Children, Toronto	
15	D2291	DNA	Dr. Chih-Ping Chen (cpc_mmh@yahoo.com)	Chen et al., 1997
16	SADC	Lymphoblast	Dr Ross Mcleod (ross@ach.ucalgary.ca), Hospital for Sick Children, Toronto	

Appendix 2

Markers used for mapping chromosome breakpoints in patients (Chapter Two)

Marker	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**HET value**	**MgCl$_2$ (mM)**	**PCR cycle**
D14S1070	F-TGGTCAAACATCTCAGAACCAG R-CAAAGTTCCAGGGCTATGTCC	244-272	0.78	1	1
D14S590	F-GTGTTGTCCTGGCGCTGT R-TAACCCCAAGGAGACACAGA	381		1.5	3
D14S1220	F-AGGGTATGAAGACAGATCTCAAGG R-ATCCCAGCCTGACTGCTG	116		1.5	3
D14S64	F-GGGCAACACAGTGAGACTCT R-TGGGATAGAAGCAACACAGA	126-136	0.77	1.5	3
[*]*D14S275*	F-TAGGAAAGCATTCCCTATTCTGGAC R-ATCAGGTTCTTGGGTACTCAATCTC	195-205	0.71	1.5	1
[*]*D14S80*	F-CATCTACCTGCCGCAA R-TAGCCAATTTATGGATACAACTT	132-156	0.84	1.5	1
D14S835	F-GACCGTACGAGGGAATTTTAA R-GTTGGAACCAAATATCACCCA	176		1.5	3
[*]*D14S740*	F-ATGCAGTTTTCTGAGACCTACC R-TGTCACTCTGCACTTTCCTG	228		2	3
[*]*D14S262*	F-GCAGTGGACTGATGCTCC R-CCATGAAACTGGTCCCG	196-206	0.58	1	1
D14S975	F-CATACACAGACACACGGAGA R-TGCCAAATAATCAGTTTTGC	168-174	0.51	1.5	1
[*]*D14S1021*	F-AGTCGTGTATCCTGGGCAT R-GCGCTGGTGTGAATCTTTA	260-276	0.64	1.5	3
[*]*D14S297*	F-ATGACAAGCATGATAAAAAGAGG R-CTGTTCAAATTTGTCACTCTCC	100	0.68	1.5	1
D14S1236	F-AGATCACTAATTTTCCTTTTCCCC R-AAATTATGTTGGTTTGAAAGGAGG	231		1.5	1
[1]*D14S730*	F-AAAAAGCCTTCACTGCAATAGC R-CCCCTCATTCTGATTTGGC	257		1.5	3
[1,*]*D14S49*	F-TCTACAAAAAGTCAGATACCT R-GAATCTTAAGTAGTTATCCCTC	167-190	0.81	2	1
[1,*]*D14S70*	F-ATCAATTTGCTAGTTTGGCA R-AGCTAATGACTTAGACACGTTGTAG	212-220	0.77	1.5	1
[1,*]*D14S599*	F-AACTTACTGGCGCAGATGAC R-ACTCAGTGCAATATTTGCTGC	90	0.86	2	4
[1,*]*D14S988*	F-TGGTGATTGGATATCACTGG R-ATGTTATGTAAGGTTTTGTTTTGTT	113-119	0.45	1.5	3
[1,*]*D14S1014*	F-AGCTATTCAGGTCAAAAAGGTC R-AATCCCTACCCTTGTGGTG	236-246	0.73	1.5	4
[1,*]*D14S888*	F-TCAAGTGCCATTTGTATTGACA R-TCTACCAATATTGTAGTACCGGAGA	190-200	0.7	2	2
[1,*]*D14S253*	F-CCAGCATAAAAAACAGACAT R-TGTGCCCGGACAGATT	141-151	0.55	1.5	1
[1,*]*D14S69*	F-AAAGCCACACTGCTAGTCAC R-TTCAGATGCCAATTAAGGGA	205-213	0.7	1.5	1

Appendix 2 continued

Markers used for mapping chromosome breakpoints in patients (Chapter Two)

Marker	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**HET value**	**$MgCl_2$ (mM)**	**PCR cycle**
[1,*]_D14S75_	F-TGTCCCCAGGTGTTGA R-CCAAGTGGCCTTGCTC	184-202	0.78	1	1
[1]_AFM205XG_[illegible]	F- TGCCATTCTCTAAGGAGCG R- CAAAATCCGACCACAGGTT	145		2	2
[1]_AFM200ZH4_	F- GTGGATGGAAATCAGGG R- CATACTGGGAATTAAAAGATGTG	230		1	3
[1]_D14S306_	F- AAAGCTACATCCAAATTAGGTAGG R- TGACAAAGAAACTAAAATGTCCC	190-210	0.85	1.5	2
[1]_D14S728_	F- TCCCACTAGACAGGGATGCT R- TATTTAGGGAGAAATCTCAATGCC	183		1	1
[1]_D14S286_	F- CATTAAGCGAAATTGGAGAA R- TTGTTTGGCCTTTGTTTATG	157-181	0.7	1.5	3
[1]_D14S738_	F- ATGACACCAGGACTTGGAAA R- TGTCAAATAGGGTGCCTCAT	270	0.75	1.5	3
[1]_D14S278_	F-GGACAAAATAATGGAATGTATAAAG R- ATCCAGGTTGCCGTGA	150-158	0.71	1.5	3
[1]_D14S600_	F- CATACTGCCGAAAGCAATCT R- TAGATGGATTGCGCCATC	150		1.5	3
_D14S872_	F-TCCAAAGGGGAAAGAGTGG TGGGAATTGGGCTTTTAGTG	100		1	3
_D14S777_	F- CAGGGAGAGGTTAAAGGACAG R- GACACAATGACAATAACTTTTGCC	167		1	1
_D14S1013_	F- GGAAACAAAATCACCAACTC R- GGCAAATGGGCAGAAT	209-241	0.84	1.5	4
*_D14S266_	F- ACAAGCCCCATATATTCATG R- AATAGACTTCCAAATCTTCAGAGTA	132-138	0.62	1.5	1
*_D14S288_	F- AGCTAGACTCTGCCATAAACA R- TGGAGACAGGAACAACACAC	189-209	0.84	1	1
*_D14S976_	F- GATACACAGCCCTGAAGTACAAAT R- GCGGTATAGACATTGTTTCCA	168-180	0.81	1.5	1
_D14S269_	F- CACATGGCATTACCAACC R- GCAACATGCTTGACAGG	213-229	0.69	1	1
*_D14S978_	F- AATACATTGGCCCCAGAG R- GGCTACATGATGGACATTCAC	228-258	0.84	1.5	1
_D14S276_	F- TGCTTTACCAAGTGCATCAC R- AGCTCAGAATCTAGGCCCT	86-98	0.77	1.5	1
_D14S52_	F- TTACTCCCTGCAAAACAAAC R- GATGAATTTCAGAAATGGAG	79-99	0.68	1.5	1
_D14S66_	F- GGCAACAGACTTGACCAATC R- CGTTCAGTAAGCAGAGAGCA	186-194	0.67	1	1
_D14S575_	F- AAGACATCTAAATTAGCAGATTTCA R- CCTGCATGACTAAAATTGAGC	302		1.5	1
_D14S592_	F- TTCCAGAGTATTTGCTTAAGAGG R- GCATTGTGGGATGAGGTATG	237	0.75	1.5	1

Appendix 2 continued

Markers used for mapping chromosome breakpoints in patients (Chapter Two)

Marker	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**HET value**	**$MgCl_2$ (mM)**	**PCR cycle**
D14S1059	F- GGTGAGGGCAGGTTGTC R- CCATTTGCGTTATCAGGC	141-155	0.60	1.5	1
**D14S63*	F- GGCCAGGTTTCAATCAGTTT R- GCCAGAGAGCCACACTGTAT	199-217	0.78	1	1
D14S251	F- AAAGGATGAACTATTGGTGC R- TTTACTTGTACCCAGTATGTNTCTG	298-318	0.83	1.5	1

*=markers used to determine parent of origin for deletion cases
1=markers used to search for microdeletions in case 16
Heterozygosity (HET) values are taken from The Genome Database (www.gdb.org), and Research Genetics (www.resgen.com)

20 pmole of primer was used per reaction.

PCR cycles:

1=94°C-4 mins, 30 cycles of 94°C-30 sec, 55°C-30 sec, 72°C-30 sec; followed by 72°C-5 mins final extension
2=94°C-4 mins, 30 cycles of 94°C-30 sec, 52°C-30 sec; followed by 72°C-5 mins final extension
3=94°C-4 mins, 30 cycles of 94°C-30 sec, 57°C-30 sec; followed by 72°C-5 mins final extension
4=94°C-4 mins, 30 cycles of 94°C-30 sec, 55°C-30 sec; followed by 72°C-5 mins final extension

Appendix 3

Markers used for uniparental disomy analysis of case 16 (Chapter Two)

Marker	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**HET value**	**$MgCl_2$ (mM)**	**PCR cycle**
D14S49	See appendix 2.2				
D14S70	See appendix 2.2				
D14S599	See appendix 2.2				
D14S988	See appendix 2.2				
D14S1014	See appendix 2.2				
D14S596	See appendix 2.2				
D14S306	See appendix 2.2				
D14S69	See appendix 2.2				
D14S52	See appendix 2.2				
D14S63	F- GGCCAGGTTTCAATCAGTTT R- GCCAGAGAGCCACACTGTAT	199-217	0.78	1	1
D14S48	F- CATAAAAGGCTTATTGGTTTG R- CAAAACAGAGAACAGAGTAG	260-275	0.82	1.5	1
D14S81	F- CAGAGAAATGAGTTGAGTATGGTT R- CAACAGAGCAGGACCCTTTC	175-209	0.84	1.5	1
D14S51	F- GATTCTGCACCCCTAAATCC R- ATGCTCAATGAACAGCCTGA	131-149	0.77	1.5	1
D14S45	F- TCCGTGAGTGATATCTTCTC R- CAGCATTACACAGGTACCAA	77-97	0.79	1.5	1
D14S272	F- GAGTTCAAGGTTACAGTAAGTNATG R- CTCTTGTCTCATAGTGCAAAGG	224-232	0.47	1.5	1
D14S1007	F- AGCTCCTATATGTCTTCACACAG R- CTCCATTCCCATACGTCC	114-132	0.77	1.5	1
D22S431	F- ATTACACTCCAGCCTGGGT R- AGATCATTGGAGAGCATTGC	249	0.67	1.5	1
D22S1265	F- CAGGACAGTTCATGGAGCTT R- ATCCTTAAAGGCTCGGCTTA	218	0.47	1.5	1
D22S1045	F- GCTAGATTTTCCCCGATGAT R- ATGTAAAGTGCTCTCAAGAGTGC	153	0.64	1.5	1
D22S417	F- CCTGGGAAGTTAAGACTGC R- TCTACCGCTTATTTCTTCCCT	183	0.86	1.5	1
D22S532	F- AGGAGGCAGAGGTTACAGTA R- GTGGTCTGAGAAGATACTTGA	185	0.63	1.5	1

20 pmole of primer was used per reaction.

PCR cycles:
1=94°C-4 mins, 30 cycles of 94°C-30 sec, 55°C-30 sec, 72°C-30 sec; followed by 72°C-5 mins final extension

Appendix 4

Markers used for mapping on CEPH YAC clones at HPE8 locus (Chapter Three)

Marker	**Primer sequence** (5' → 3')	**$MgCl_2$ (mM)**	**PCR Product size (bp)**	**PCR cycle**
D14S1236	F-AGATCACTAATTTTCCTTTTCCCC R- AAATTATGTTGGTTTGAAAGGAGG	1.5	231	1
D14S70	F-ATCAATTTGCTAGTTTGGCA R-AGCTAATGACTTAGACACGTTGTAG	1.5	212-220	1
D14S49	F-TCTACAAAAAGTCAGATACCT R-GAATCTTAAGTAGTTATCCCTC	2	169	1
D14S888	F-TCAAGTGCCATTTGTATTGACA R-TCTACCAATATTGTAGTACCGGAGA	2	199	2
D14S1014	F-AGCTATTCAGGTCAAAAAGGTC R-AATCCCTACCCTTGTGGTG	1.5	236	3
D14S599	F-AACTTACTGGCGCAGATGAC R-ACTCAGTGCAATATTTGCTGC	2	90	3
D14S730	F-AAAAAGCCTTCACTGCAATAGC R-CCCCTCATTCTGATTTGGC	1.5	257	4
D14S723	F-CCCAACTAGGCTATTAGCAACA R-GCCTCAATGTTTTTAAAGGAGTT	1.5	151	4
D14S760	F-TTGGTTGGCATTTTAGCCTC R-AGTTTCAGGAGCATACTGATAAACA	1.5	260	4
D14S596	F-AACTTACTGGCGCAGATGAC R-TGGAATAAGGCCTGACTCAG	1.5	104	4

PCR cycles:

1= 94°C - 3mins, 30 cycles of 94°C - 30 sec, 55°C - 30 sec and 72°C - 30 sec; followed by 72°C-5 mins final extension
2=94°C - 3 mins, 30 cycles of 94°C-30 sec and 52°C-30 sec; followed by 72°C-5 mins final extension
3= 94°C - 3 mins, 30 cycles of 94°C-30 sec and 55°C-30 sec; followed by 72°C-5 mins final extension
4= 94°C - 3 mins, 30 cycles of 94°C-30 sec and 57°C-30 sec; followed by 72°C-5 mins final extension

Appendix 5

BAC clones used for construction of physical map (Chapter Three)

R159D23	R808K16	R797P01	R130E15	R583C14
R530D13	R528E14	R64M09	R32L09	R457P22
R340L03	R797J06	R388F21	R271C19	R743N09
R431H16	R242O14	R79O05	R434022	R512B22
R356O09	R320M16	R811D01	R267E16	R49N16
R349I04	R187E13	R198A09	R433O19	R320N05
R169C16	R142C01	R510I24	R33H24	R512M21
R1078I14	R501E21	R759P18	R256E06	R379L15
R434N21	R138H18	R973N08	R436G18	R729P12
R702A1	R422L13	R633C12	R530D13	R702A01
R21G05	R151B16	R247E17	R546B03	R261B19
R808G18	R69D20	R158E19	R260C03	R399J10
R114L12	R273H04	R209I01	R744M19	R359C03
R134E15	R33H21	R133A04	R378O06	R759F09
R372C16	R141J18	R501H04	R98M16	R578G22
R808P18	R138E07	R813F13	R715N05	R336P22
R589O07	R679N08	R404G21	R412G17	R311O13
R588D07	R642D01	R484O07	R40D24	R332N19
R898B23	R748E05	R748O05	R519L11	R645L20

Appendix 6

PCR primer pairs used to make probes for screening BAC filters (Chapter Three)

Primer	BAC clone	Primer sequence (5' → 3')	Product Size (bp)	$MgCl_2$ (mM)
C2008K18L	C2008K18	F-AAAAACTGGCTGTCCCTCCT R-TGCAGTCTGTTGCCCAGTAG	603	2
C2008K18R	C2008K18	F-CTGGGAGGAACTGAAGGACA R-TAGGCAAGGCATCCAAAGAC	599	2
R1078L14R	R1078L14	F-CAGTAAGAATGGGCCTTCCA R-GCACTGAGCAAAGCTTGACA	593	2
R422L13R	R422L13	F-TGGAAAAATAAGGATAAAGATAGTTTG R-TTTTCCTCTGGATCTTCTGCTC	333	2
R422L13L	R422L13	F-GAGGCCTCAGAATCATGGTG R-TTGCAGGACACAAAATCAAAA	374	1.5
R1078L14L	R1078I14	F-TCGTGTGGAAAGAAATGCTG R-CATTGCACTGGTCCTGAATG	599	1.5
R434O22T7	R434O22	F-CCAGTAAGCCCAACAGCTTC R-TGGACCAAGGTCAGTGGAAC	533	2
R436G18SP6	R436G18	F-AGACAAGTTATTCACCTCCTCAGA R-AAGTAGGAGAGCCCCCAAAA	332	2
R1078L14	R1078L14	F-GGCTTAAGGAAGGGGAAACA R-AATGGGCTTCTCCAAAGGAT	476	2
R702A1SP6	R702A1	F-GGCACAAGGAAGACTGGAGA R-AGGGCCAATGTAACAAGGAA	600	2
R379L15T7	R379L15	F-GGAGAAGTTACGTGGCATCC R-TGAAGGAGCATGCTTTACGA	606	2
R33H24T7	R33H24	F-GGAAGGCGAACTACTGCTGA R-CACAGGCCTACATGGTGTTG	607	2
R702A1T7	R702A1	F-CAAATGAGCATTGGCTTCAA R-CAACTCGTGAATGCAAAGGA	549	2
R33H24SP6	R33H24	F-GGCTTAAGGAAGGGGAAACA R-AATGGGCTTCTCCAAAGGAT	603	2
260C35SP6	R260C35	F-GGTCAGGATGCAGTGTCCTT R-TTTCCCCACTAAAGTCCTGGT	541	2
379L15SP6	R379L15	F-AAAAAGTTGCAGGCCCTCTT R-TCTGCCATGGAAAAGCTAATG	548	2
457P22T7	R457P22	F-CTGGTAGTCCCACAGACAAATG R-TGCAGAAGTGGAGAATTTGC	534	2
261B19T7	R261B19	F-CACCAAGCTTCTGTAGTTTT R-TGGCCAAAACTTAATTTCTAC	244	2
729P12SP6	R729P12	F-AGAAGTGTTTTGGACTCCTTGG R-CAAAAGTACACTCAGAACTCCACTG	367	1.5
359C3T7	R359C3	F-AGCAGTCCACTATGAGGCAGA R-GCTCCTAAGTCCCTGGCCTA	317	1.5
546B3SP6	R546B3	F-TGTTAGTGGTCTCTGCCACAA R-ATTGTGCCCTGCCAGAGAT	251	1.5
578G22SP6	R578G22	F-GGTCTTGCAAACACCCAGTT R-TGTGTAACCTGAATCTAAGCAACC	478	1.5
358C3SP6	R358C3	F-TGCATAGACAAATGACTAGTATTCAGA R-ATTTCCCTAATAAGTAGTGATGCT	180	1.5

The PCR cycle was: 94°C-3 mins, 30 cycles of 94°C-30 sec, 55°C-30 sec and 72°C-30 sec; followed by 72°C-5 mins final extension

Appendix 7

Primers used for obtaining cDNA by RACE (Chapter Three)

Primer name	**Gene**	**Primer sequence** (5' → 3')	**Type of RACE**	**T m**	**PCR cycle**
GS31LMCR-R	NPAS3	TGCCTTGTCGAGCTGGCTGGTAA	5'	70	1
GS31LMCRN-R	NPAS3	GCTGGCTGGTAATGGCTGCAGAG	5' Nested	69	1
GS26LMCR-R	GS26LMCR	GCCCAACTGGTGCAGTCGTCGTA	5'	70	1
GS20LMCR-R	GS20LMCR	GGCACCCAAGGACCCACTCAAAT	5'	69	1
GS20LMCRN-R	GS20LMCR	CCGCTAAAACAGGCCACGTTAGCA	5'Nested	70	1
N26711-R	N26711	GGGCTGAGTTGGGTGGACCACTT	5'	70	1
T75473-R	T75473	TGACATGTGTGCATGGAGCTTGGA	5'	70	1
R26150-R	C14ORF11	TGCATCAACCCAGGCCTGATCTC	5'	70	1
R26150N-R	C14ORF11	CACAAGCAATTTGCGGCGTCTCT	5'Nested	70	1
H41238-R	H41238	GGATTGTGGTCCCTGCCCTCAAG	5'	70	1
H41238N-R	H41238	GGGCTCTCCCTGGGTCTCCAGCTT	5'Nested	70	1
N67117-R	C14ORF10	TCAAGGAGCATCAATTAAACAGCTTGG	5'	67	2
N67117N-R	C14ORF10	GGAGCATCAATTAAACAGCTTGGA	5'Nested	64	2
AA609891-R	AA609891	CAGGCTGGCTTTAACTGTTCATACACC	5'	67	2
GS6RMCR-R	C14ORF19	TTGGGCCATGTTCCGAGTGAGAA	5'	70	1
GS6RMCRN-R	C14ORF19	TCACTGGAGGCCTCTCCTGGCTA	5'Nested	70	1
GS13LMCR-R	GS13LMCR	TACCTCGGCCACAGGTCCTCCTC	5'	70	1
GS13LMCRN-R	GS13LMCR	CGGCCACAGGTCCTCCTCTCAG	5'Nested	70	1
H05452-R	H05452	CATCCCCACAATGGAAACCAGCA	5'	70	1
H05452N-R	H05452	GCGATGCAGCTAACACGTCAGTC	5'Nested	70	1
NS0090-R	NS0090	GGATCTATCTAGCAAAAGCCGAGA	5'	62	2
NS0090N-R	NS0090	TCACACTAAGGATGCTTCACCA	5'Nested	61	2
GS8RMCR-L	C14ORF10	CGCTATGGAACAGCTACCATGACC	3'	69	3
GS6RMCR-L	C14ORF19	TCGGAACATGGCCCAAGTCAAAG	3'	69	3
GS6RMCRN-L	C14ORF19	GAGTTCGGGAGTGGGATGACTGGA	3'Nested	70	3
GS13LMCR-L	GS13LMCR	AGGAGGACCTGTGGCCGAGGTAA	3'	69	3
GS13LMCRN-L	GS13LMCR	AGCGAGGCCCACACCACTGGTAA	3'Nested	61	4
GS31LMCR-L	NPAS3	ACGGAGCCAGCTCAGCATCTTCC	3'	70	3
GS26LMCR-L	GS26LMCR	CGACGACTGCACCAGTTGGGCTA	3'	71	3
GS26LMCRN-L	GS26LMCR	CGACTGCACCAGTTGGGCTACAC	3'Nested	68	3
GS20LMCR-L	GS20LMCR	CTCGCCAGTCTTCCTCAGGTCCA	3'	69	3
GS20LMCRN-L	GS20LMCR	TGGGTGCCATGGACAGAAATCTTG	3'Nested	69	3
N26711-L	N26711	AACCACTGTGCCAGGCCCATAGC	3'	70	3
T75473-L	T75473	CAGGCCAGAGCTTGAGGTGAACCA	3'	71	3
R26150-L	C14ORF11	AACCGGCTTCCGGATGACTACGA	3'	70	3
R26150N-L	C14ORF11	TTGTCCTGCCTGCATGACCACAC	3'Nested	70	3
H41238-L	H41238	TTGAGGGCAGGGACCACAATCCT	3'	70	3
H05452-L	H05452	TGCTGGTTTCCATTGTGGGGATG	3'	70	3
N67117-L	C14ORF10	TCCAAGCTGTTTAATTGATGCTCCT	3'	64	4
N67117N-L	C14ORF10	CAGCTTTGGGTTTGAAACTATCC	3'Nested	61	4
NS0090-L	NS0090	TCAGATCTCGGCTTTTGCTAGA	3'	61	4
AA609891-L	AA609891	AACGGATCCCACATTGTTTCTG	3'	63	4

Appendix 7 continued

PCR cycles:
1=94°C-1 min, 5 cycles of 94°C-20 sec, 71°C-4 min, 5 cycles of 94°C-20 sec, 69°C-4 min, 25 cycles of 94°C-10 sec, 67°C-4 min; followed by 72°C-5 mins final extension
2=94°C-1 min, 35 cycles of 94°C-20 sec, 61°C-20 sec, 68°C-4 min; followed by 72°C-5 mins final extension
3=94°C-1 min, 5 cycles of 94°C-20 sec, 71°C-2 min, 5 cycles of 94°C-20 sec, 69°C-2 min, 25 cycles of 94°C-10 sec, 67°C-2 min; followed by 72°C-5 mins final extension
4=94°C-1 min, 35 cycles of 94°C-20 sec, 61°C-20 sec, 68°C-2 min; followed by 72°C-5 mins final extension

Appendix 8

PCR primers used for RT-PCR analyses (Chapter three)

Primer name	Gene	Primer sequence (5'→ 3')	PCR product size (bp)	$MgCl_2$ (mM)
*Em1	GS9RMCR	F-TGTATGCCACACTGCACTCC R-TTGTTTGCACCACTGTGTCA	207	1.5
**Em2	PSMA6	F-TGATGCTCACCTTGTTGCTC R-CACTTCGGTGGTAGGAGGAG	162	1.5
**Em3	NFKBIA	F-ATCAGCCCTCATTTTGTTGC R-ACAGGATACCACTGGGGTCA	187	1.5
**Em4	INSM2	F-TTAATGGGGTTTCACCCAAA R-AGTTGCCGCTGAAAAGAAAA	166	1.5
**Em5	FLJ12926 fis	F-TTGACAATTTTGGCTTTGGA R-GGGAAGCACTTTCTGAACCA	201	1.5
**Em6	FLJ23250 fis	F-AACAAGGTTGGGACAAGCAC R-AATTTCCCCAAAGCGTAACC	184	1.5
**Em7	Hypothetical protein MGC11296	F-AGGCCTGATGGAAGCAAATC R-CTCTCATGGCACTTACATTTGG	138	1.5
**Em8	MBIP	F-GCATTGAAAATAGCATTGTTGG R-CAAAATAAATCTTCCTTCCTGT	134	1.5
**Em10	NKX2.9	F-GGTCTCCTGGAACTGGTGAG R-ACGTCCCTCGTGTGTGCT	198	1.5
**Em11	Hypothetical protein MGC15427	F-AATCTTGCTTTTCCCACACG R-CCTAGGGGTAGGGAGGGGTA	212	1.5
**Em13	SLC25A21	F-AACTTTGGGACGACATGGAG R-ACTGGTTGAGGCCCTTGAAT	201	1.5
*Em14	GS4RMCR	F-TGCTGGCTGCAGGTATTAAG R-GGAAAAAGGGAGGGTGTTACTT	196	1.5
*Em15	GS5RMCR	F-TTCGTTCAGTGAGAAGCTGGT R-TGCTGGTGGAACACATCAAT	202	1.5
*Em16	GS7RMCR	F-TCCCCTCCTTTTCTTTTTCC R-AATGCTTGGCCTTCGAGATA	189	1.5
*Em17	GS24LMCR	F-TTTTCCTCTTCTGGTTTTGGAA R-GCTGAGAGTATGTTGCATTGG	219	1.5
*Em18	GS2RMCR	F-ATTCTGGGCTCCTGAAAGTG R-CATTGCCTCCCAATTTCTCT	221	1.5
*Em19	N26711	F-TGAGGCCATATTTACCAATCC R-TCCAGTTCTAGGAAGCCCAGT	218	1.5
*Em20	T75473	F-ATGCACACATGTCAGATGGA R-CCATCAGGTCCCTCATTTTG	191	1.5
*Em21	H41238	F-GGCCTGAGAACTCAGAGAGC R-TGCCCTCAAGATGCTTACCT	203	1.5
*Em22	H05452	F-TCTTCCATATGTCTCACAGGTCA R-GTGAAGTTGAAAGGCCAGGA	185	1.5
*Em23	N50090	F-TTCAGATCTCGGCTTTTGCT R-TCTCTGTCCCACCATAAAATGA	194	1.5
*Em24	AA609891	F-CATCTTTGTCGAATAACCATTCC R-TTCGAAACTAGAAGTAACTCAGAACG	145	1.5
*Em25	GS26LMCR	F-TGGGACCAGAGAGGCTAGAA R-TATTGTGCTGCAGAGCCAAC	152	1.5

*=genes mapping within HPE8
**=genes mapping within D14S1014 and AFM205XG5
50 ng of each primer was used per 20 µl PCR reaction; The PCR cycle was 94°C-5 min, and 30 cycles of 94°C-30 sec, 57°C-30 sec; followed by 72°C-5 mins final extension

Appendix 9

PCR primers used to make probes for screening λ TriplEx cDNA library (Chapter Four)

Probe	**Gene**	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**$MgCl_2$ (mM)**
RD5	C14ORF11	F-GGCATGAATCATACAAAACTCAA R-TCCAGAACATTTTAATTCTACAGAGC	567	2
RD7	C14ORF10	F-TGAAAAGGCTGGAGCAAAGT R-TGAGCATGCCATTGTGATCT	500	2
RD8	NPAS3	F-TCTTGGGGAGCAGAAGGTAA R-AGATTCTGCCCTCAGCAATG	437	2
RD9	C14ORF19	F-GCGTAGGAACCGAAAGAAGA R-TTGAGGTGAACCCAGAGCTT	499	2

PCR cycle:
94°C=3 min, 30 cycles of 94°C-30 sec, 58°C-30 sec, 72°C-30 sec; followed by 72°C-5 mins final extension

Appendix 10

PCR primers used to make EGFP-N1 constructs (Chapter Four)

Gene/plasmid	**Primer sequence** (5' → 3')	**PCR cycle**
NPAS3	F-AGGCATCCATCATTCGACTT R-CGTGGTGTAGACCCTCTGC	1
C14ORF11	F-ATGAACCGGCTTCCGGATGACTAC R-TGGACCCGCCTTTTCTTCCTTGTT	2
SNX6	F-ATGACGAAGGAAGAATTCACAAAGATG R-TGTCTCCATTTAACACTGCCAGGC	3
C14ORF11 engineered	M13 forward R-**GGATCC**GGACCCGCCTTTTCTTC	4
NPAS3 engineered	F-**AGATCT**ATGAACCACATTTTGCAGTCCCTGGATG R-**GGATCC**GAGGACCGAGTCGGGAATGGC	5
EGFP-N1	F-GTAACAACTCCGCCCCATT R-GCTGAACTTGTGGCCGTTTA	6

M13 forward sequence = CGCCAGGGTTTTTCCAGTCACG

PCR cycle:

1= denaturation for 94°C-3 min, 35 cycles of 94°C-30 sec, 60°C-30 sec and 68°C-150 sec, followed by 72°C-5 min
2= denaturation at 94°C-3 min, 36 cycles of 94°C-30 sec, 55°C-30 sec and 68°C-75 sec, followed by 72°C-5 min
3= denaturation at 94°C-3 min, 35 cycles of 94°C-30 sec, 54°C-30 sec and 68°C- 60 sec, followed by 72°C-5 min
4= denaturation at 94°C-3 mins, 35 cycles of 94°C-30 sec, 56°C-30 sec, 72°C-90 sec, followed by 72°C-5 mins
5= denaturation at 94°C-3 min, 6 cycles of 94°C-30 sec, 53°C-30 sec and 72°C-90sec, 25 cycles of 94°C-30 sec, 64°C-30 sec and 72°C-90sec, followed by a final extension of 72°C-5 mins
6= denaturation at 94°C-3 mins, 30 cycles of 94°C-30 sec, 57°C-30 sec, 72°C varying from 60 sec to 90 sec, followed by 72°C-5 min

Appendix 11

Assembled cDNA sequences for HPE candidate genes (Chapter Four)

NPAS3 cDNA (large) (3392 bp) [neuronal PAS3 transcript variant 1] Genbank accession number AY157302

acgccccccacccgggagggggggagagaggcaaaaagtaagagaggaaaaaaaatagcaggaagatggcgcccaccaagcccagctttcagcagg
atccttccaggcgagaacgtttacaagcattgagaaaggagaaatcccgagatgctgctcgctcccgccggggaaaagaaaactttgagttctatgaattgg
ccaagttgttgcctcttcctgcagccattaccagccagctcgacaaggcatccatcattcgacttacaattagctatctgaaaatgagggactttgctaaccag
ggggaccctccgtggaacttgcgaatggaaggccctccacctaacacatcagtaaaaggtgcacagcgaaggagaagccccagtgcactagccattgaa
gtatttgaagcacatttgggaagccacattttgcagtccctggatggctttgtatttgcactaaatcaggaaggaaaatttttgtacatttccgaaacagtctccat
ctacctaggcctctcacaagtggagctgacaggcagcagtgtctttgactatgtccaccccggagatcacgtggagatggctgagcagctgggcatgaag
ctccccccctgggcggggtctcctgtcacagggcactgctgaggacggagccagctcagcatcttcctcctctcagtcggagacccccgagccagtggagt
caaccagccccagtctgctaaccactgacaacactcttgagcgttccttttcatccgaatgaaatctactctgaccaaacgcggtgtgcacatcaaatcatca
ggatataaggtgattcacataacaggccggctacgcctgagagtgtcgctgtcccacgggaggaccgtccccagccaaatcatgggtctcgtggttgttgc
gcatgccttgcctccccctacgatcaatgaagtcagaattgactgccatatgttcgtcactcgagtaaatatggacctcaatatcatttactgtgaaaataggatt
agtgattatatggatctgacccctgtagatatcgtagggaagagatgctaccacttcatccatgctgaagacgtggagggcatcaggcacagtcacttggact
tgctgaataagggtcagtgtgtgacaaagtactatcgctggatgcagaagaacggaggatatatttggatacagtccagtgccaccatagctattaatgccaa
gaatgcaaatgaaaagaatatcatctgggtgaattaccttcttagcaatcctgagtacaaggacacacccatggacatcgcacagctcccccatctgccgga
gaaaacttccgaatcctcggagacatccgactctgagtcagactctaaagacacctcaggtattacagaggacaacgagaactccaagtccgacgagaag
gggaaccagtccgagaacagcgaagacccggagcccgaccggaagaagtcgggcaacgcgtgtgacaacgacatgaactgcaacgacgacggcca
cagctccagtaacccggacagccgcgacagcgacgacagcttcgagcactcggactttgagaaccccaaggcgggcgaggacggcttcggtgctctgg
gcccgatgcagatcaaggtggagcgctacgtggagagcgagtcggacctgcggctgcagaactgcgagtcactcacgtccgacagcgccaaggactc
ggacagcgcaggcgaggcgggcgcgcaggcctccagcaagcaccagaagcgcaagaaaaggcggaaacggcaaaagggcggcagcgccagccg
ccggcgcctttccagcgcgtcgagcccaggcggcctggacgcgggcctggtggagcccccgcggctgctgtcctcccccaacagtgcctcggtgctca
agatcaagacggagatctcagaacccatcaatttcgacaatgacagcagcatctggaactacccgcccaaccgggagatctccaggaacgagtcccccta
cagcatgaccaagcccccagctctgagcacttcccgtccccgcagggcggcggcggtgggggtggcggtggcggggggctgcacgtggccattccc
gactcggtcctcaccccgcccggcgccgacggcgcggccgcccgcaagactcagttcggcgcctcggccaccgcggccctggcccccgtcgcctccg
acccgctgtcacccccgctctcggcgtccccgcgggacaagcaccccgggaacggcggcgggggcgggggcgggggcggcggcgcggggggcg
gcggccccagcgcgtccaactccttgctgtacactggggacctggaggcgctgcagaggttgcaggcgggcaacgtcgtgctcccgctggtgcacagg
gtgaccgggaccctggccgccaccagcacggccgcgcagagggtctacaccacgggcaccatccgctacgcgcccgccgaggtgaccctggccatg
cagagcaacctgctgcccaacgcgcacgctgttaacttcgtggacgttaacagccccggctttggcctcgaccccaagacgcccatggagatgctctacca
ccacgtgcaccggctcaacatgtcaggaccgttcggcggcgcagtgagcgcagctagcctgacgcagatgcccgccggcaacgtgttcaccacggccg
agggactcttctccacgctgcccttccccgtctacagcaacggcatccacgcggcacagactctggagcgcaaggaggactgaggcgccgcccgtcctg
ggcccggccaggccccgcttggaggaggcatcgtcggcattttcgtttagacctttaattctagcactttgaattcgagcaggtcagcgtcttctctcgccacg
acggtccccattccacccccctctttctttcacctgacttattctttcgtgtaaagatatgtttatttttgccttcagagggtcagacgaccagttgcctgccgttttgt
cttcttctaaggtgtgtgttgggttgttttgctttcctttgcatctttattaagatgtctttcatgtgtatatgcctctgccatagaatactcagtcttgtggtcaagagag
ttctcaagtgacaaccattggggtttcttcataaagatcttgatatgatcaagatggaaagagacaagcataaacaatgtgccctgtttgactaagtcaaatgaa
atagggtggtttttgtttctgttcctaattcctttaaaaaataggggggaatagtattttagaattttatgcagaatctaattctctttttacggttaagattctaagattatc
ttacttgcacataataaatttgggttcttaaacttaatttctggcctgt

NPAS3 cDNA (small) (2530 bp) [neuronal PAS3 transcript variant 2] Genbank accession number AY157303

gattttgtcacctgagactgctgttactctgaaaatgatagcattaacttgtctaattctgagaaagacatttgttagctaatgtagactattaatagagcaaacata
aaactggttgggtattagcctagcaagtatctgaccagagcattatggcattacttagattacatagatattcagtaagatatttcaggtgtgatttctgcagtgaa
tacagtgtttacaccagaaaacaatgtcaggaatcctataatgtaatagggaatattattcaccatagtatattaaaagagaacggagccttgctgctgaaccaa
ttttgatctcttcctttggtcttccttccttccttgtttctctctttcttattatttcataatgaaaaaaaagaatgctagactcaaatatttacagtgctttcagaattcaata
aaatgcatacataaaaatgattttaatagtaccttgatactgaaaatattatctagctacttatcattgtggtctgatttttaaatttatttctcatcagacacaatacag
agcaattcagagcaggtagaattcttctgtcagatgctcacacacaacttcagttgtgccagattctgccctcagcaatgtgagcaaacttcaaacagaaatgc
acatgaatagttgaaggaaaatgctggatcccacctgttgcctcttctgggaagtaatatgtctataatgtttataaattttaataacccatacagcgctttatctttc
aatattaaatttatattttcctaaagtttgacttctttgggcaatatattttattctctagtatcttagaaaatcaatctactaatgtatgtccattaataaataagtatgctc
ttatcaagcagttatcagcacatttataaaattacattgaaatcttcattttaaatatgatgttttcttcatctaactgaacatgctgtgatctttattgagggctggaaa

catatttgttgattctgacaaggtgaattgtattcgtgtttaccttctgctccccaagattataattgaaatgaatttttcaaaatagttccttttgatgaagatcccata
ctgatgaaagttaaatgaacagaagtattagtttgaaatagatacacgaatgtaagcccaacttataaattgattctagttgataatatcaagctacgcctgtaaa
ccttctgttaggtcttagaaacaataaatacatctcaatgcatggacaattgaaggaaaaaaaccttaacaagtaaaacacaatttttactcattggacttgagttt
acaaaatgtaaggccagagttagagtttatggctttagtggacattttctggaactcattttactgtcactctctagttttgaaatagtttttggtagaattttggtggg
aaaaaaggaaatacaaagaaagcagtatttgaatgatgacagagtcacatattctaaaccacacattctcactcctttgatttcagtttacaagcattgagaaag
gagaaatcccgagatgctgctcgctcccgccggggaaaagaaaactttgagttctatgaattggccaagttgttgcctcttcctgcagccattaccagccagc
tcgacaaggcatccatcattcgacttacaattagctatctgaaaatgagggactttgctaaccagggggaccctccgtggaacttgcgaatggaaggccctc
cacctaacacatcagtaaaaggtgcacagcgaaggagaagccccagtgcactagccattgaagtatttgaagcacatttgggaagccacattttgcagtcc
ctggatggctttgtatttgcactaaatcaggaaggaaaatttttgtacatttccgaaacagtctccatctacctaggcctctcacaagtggagctgacaggcagc
agtgtctttgactatgtccaccccggagatcacgtggagatggctgagcagctgggcatgaagctccccccctgggcggggtctcctgtcacagggcactgc
tgaggacggagccagctcagcatcttcctcctctcagtcggagacccccgagccagtggtttgcttcccaccagcgagtgatcagtttcttctgtaggacccc
tgtcaggaagaacatgcctggatcctacttcagagagaactcagccagggcaaggggctcttgtggctttttgggtgggtgtgagttcatagaggggaaaat
caatactgtaccaggatttagacaatttagcaaaatgcatgaggcagtactttggatgatgaataactacagctgcatggctaacccattggttgtgtaaagtat
gggttatcatgttttgtctaaattttcaggcaaaatctagcctgttttgtggcattaggatcatcattatagtagtatgatttttatgctattatacttacgctatgtcccg
aatgtttttaaggccagcaaaatggaaacagtagagtacttgcctgaatgcaaataaacttcagtttcaactt

C14ORF10 cDNA (large) (1843 bp) [chromosome 14 open reading frame 10 transcript variant 1] Genbank accession number AY157304

ctaatgaactcctgggtcacgtggtgcgcacgcgccaaaggcacacgataaaggaaaacgcgggcttatcgtggcctttaaaatccggcttggtgagcttg
ggtcgcctctgaaggagaaccattttccatctctttcatagtttttcccccagtcagcgtggtagcggtattctccgcggcagtgacagtaattgtttttgcctctt
tagccaagacttccgccctcgatcaagatggtggttggacggccttcctaacctttacggggcctggcggtgctgacgcctgagctggtaggggtggagca
ggtaggaaacagcaaatgcagaagctgctgcgcggaagtcggccatggactggaaagaagttcttcgtcggcgcctagcgacgcccaacacctgtccaa
acaaaaaaaaaagtgaacaagaattaaaagatgaagaaatggatttatttacaaaatattactccgaatggaaaggaggtagaaaaaacacaaatgaattcta
taagaccattccccggttttattataggctgcctgctgaagatgaagtcttactacagaaattaagagaggaatcaagagctgtctttctacaaagaaaaagca
gagaactgttagataatgaagaattacagaacttatggtttttgctggacaaacaccagacaccacctatgattggagaggaagcgatgatcaattacgaaaa
ctttttgaaggttggtgaaaaggctggagcaaagtgcaagcaatttttcacagcaaaagtctttgctaaactccttcatacagattcatatggaagaatttccatc
atgcagttctttaattatgtcatgagaaaagtttggcttcatcaaacaagaataggactcagtttatatgatgtcgctgggcaggggtaccttcgggaatctgattt
agaaaactacatattggaacttatccctacgttgccacaattagatggtctggaaaaatctttctactccttttatgtttgtacagcagttaggaagttcttcttcttttt
agatcctttaagaacaggaaagataaaaattcaagatattttagcatgcagcttcctagatgatttattggagctaagggatgaggaactgtccaaggagagtc
aagaaacaaattggttttctgctccttctgccctaagagtttatggccagtacttgaatcttgataaagatcacaatggcatgctcagtaaagaagaactctcacg
ctatggaacagctaccatgaccaatgtcttcttagaccgtgttttccaggagtgtctcacttatgatggagaaatggactataagacctacttggactttgtccttg
cattagaaaacagaaaggaacctgcagctctacaatatattttcaaactgcttgatattgagaacaaaggatacctgaatgtcttttcacttaattatttctttaggg
ccatacaggaactaatgaaaatccatggacaagatcctgtttcatttcaagatgtcaaggatgaaatctttgacatggtaaaaccaaaggatcctttgaaaatct
ctcttcaggatttaatcaacagtaatcaaggagacacagtaaccaccattctaatcgatttgaatggcttctggacttacgagaacagagaggctcttgttgcaa
atgacagtgaaaactctgcagaccttgatgatacatgatctctgaaagactagactgtcttatattatgagatacttgaatgctgcatgtaaagcctttaaagcaa
aatcctcagaaatggtctaaataaaacacttgatatgcctagagaacacaaaaaaaaaa

C14ORF10 cDNA (small) (1342 bp) [chromosome 14 open reading frame 10 transcript variant 2] Genbank accession number AY157305

ctaatgaactcctgggtcacgtggtgcgcacgcgccaaaggcacacgataaaggaaaacgcgggcttatcgtggcctttaaaatccggcttggtgagcttg
ggtcgcctctgaaggagaaccattttccatctctttcatagtttttcccccagtcagcgtggtagcggtattctccgcggcagtgacagtaattgtttttgcctctt
tagccaagacttccgccctcgatcaagatggtggttggacggccttcctaacctttacggggcctggcggtgctgacgcctgagctggtaggggtggagca
ggtaggaaacagcaaatgcagaagctgctgcgcggaagtcggccatggactggaaagaagttcttcgtcggcgcctagcgacgcccaacacctgtccaa
acaaaaaaaaaagtgaacaagaattaaaagatgaagaaatggatttatttacaaaatattactccgaatggaaaggaggtagaaaaaacacaaatgaattcta
taagaccattccccggttttattataggctgcctgctgaagatgaagtcttactacagaaattaagagaggaatcaagagctgtctttctacaaagaaaaagca
gagaactgttagataatgaagaattacagaacttatggtttttgctggacaaacaccagacaccacctatgattggagaggaagcgatgatcaattacgaaaa
ctttttgaaggttggtgaaaaggctggagcaaagtgcaagcaatttttcacagcaaaagtctttgctaaactccttcatacagattcatatggaagaatttccatc
atgcagttctttaattatgtcatgagaaaaggtgattcatttttactgtgcttgaactccttatctgggaatcaagtaggtttttggtctgtgcttttataagtatttcttct
atgtgaaatcactctaagtttgaatagtagaatttcttatcttttgaccaagatgtgttacataaacttgggttagcataaaaaaacaaaaatttttaaaccctcatgt
gtcttaaagataatagtagactagaaatattaagagctgaaccaagagttttaatgtatggtaatttgaaataattttaatttactgataattttattttaccagtatatt
acaagttgagttgactgaaaggatagtttcaaacccaaagctgtaatatcagtttcaaggagcatcaattaaacagcttggacttatgattcctacaagtgccat
aataatctgtgaaagcctattgtatacaatctttaataaattggttgctgtaattaacctcagaaataaaaaaaaaa

C14ORF11 cDNA (1348 bp) [chromosome 14 open reading frame 11] Genbank accession AY157301

aaatttagcggccgcgaattcgccctttccgagatctggacgagcttttttttttttttctcgggaagcgcgccattgtgttggtacccgggaattcggcacaagg
gggcgctttagggactggacttgcagtgtaaacagagacgctgcaaattgcttgtggacggtgtaggccgctgcaggccaccatgaaccggcttccggat
gactacgacccctacgcggttgaagagcctagcgacgaggagccggctttgagcagctctgaggatgaagtggatgtgcttttacatggaactcctgacca
aaaacgaaaactcatcagagaatgtcttaccggagaaagtgaatcatctagtgaagatgaatttgaaaaggagatggaagctgaattaaattctaccatgaaa
acaatggaggacaagttatcctctctgggaactggatcttcctcaggaaatggaaaagttgcaacagctccgacaaggtactacgatgatatatattttgattct
gattccgaggatgaagacagagcagtacaggtgaccaagaaaaaaaagaagaaacaacacaagattccaacaaatgacgaattactgtatgatcctgaaa
aagataacagagatcaggcctgggttgatgcacagagaaggggttaccatggtttgggaccacagagatcacgtcaacaacagcctgttccaaatagtgat
gctggtcttgaattgtcctgcctgcatgaccacactcttgccttgattgccaaaggcatgaatcatacaaaaactcaatatagagcaatgtttgtaatgaattgttc
tattaacaaagaggaggttctaagatataaagcctcagagaacaaggaagaaaaggcgggtccataagaagatgaggtctaaccgggaagatgctgccg
agaaggcagagacagatgtggaagaaatctatcacccagtcatgtgcactgaatgttccactgaagtggcagtctacgacaaggatgaagtctttcatttttc
aatgttttagcaagccattcctaaacagcccaactggcatttaattacccaatactgtatataaggcaaatatggacagttactttcctcttgcctgttcatatccttc
agtgacattgaggaagcagtgtttctctttttaaaggagaatagttgtcaaccttcattcatctcttacatctctcaccctctcctttttttttttctttgattttccccctta t
tgatgggactgatattcattctgtttttgatgaacatttggaaactgtcgggctttttattaaagctctgtagaataaaatgttctggaattataagc

C14ORF19 partial cDNA (600 bp) [chromosome 14 open reading frame 19] Genbank accession AY168620

atcagcttagaagagattgagagcattgaccaggaaataaagatcctgagagaaagagactcttcaagagaggcatcaacttctaattcatctcgccaggag
aggcctccagtgaaacccttcattctcactcggaacatggcccaagtcaaagtatttggagctggttatccaagtctggcaactatgacggtgagtgactggt
ataagcaacattggaaatatggagcattaccagatcagggaatagccaaggcaacaccagaggaattcagaaaagcagctcagcaacaggaaggtcaag
aagaaaaggaggaagaggatgatgaacaaacacttcggcttcagagttcgggagtgggatgactggaagggcatccatcctagggactatggcagccaa
cagagcatgggctgatcttcccacaaatgagaggactgaagggtgcacacctcccacgccaaggaaaaccatgcagtcctcatctccctgggcttccgctt
cagctgtgtacaacgaaggcaaagatgctaaattttgctttgctttcaataaagtgtccagtgattaaaaaaattagaaa

Appendix 12

Protein sequences of the holoprosencephaly candidate genes (Chapter Four)

NPAS3 (901 amino acids) (large) [neuronal PAS3 protein variant 1] Genbank accession AAO17043

MAPTKPSFQQDPSRRERLQALRKEKSRDAARSRRGKENFEFYELAKLLPLPAAITSQLDKASIIRLTI
SYLKMRDFANQGDPPWNLRMEGPPPNTSVKGAQRRRSPSALAIEVFEAHLGSHILQSLDGFVFAL
NQEGKFLYISETVSIYLGLSQVELTGSSVFDYVHPGDHVEMAEQLGMKLPPGRGLLSQGTAEDGA
SSASSSSQSETPEPVESTSPSLLTTDNTLERSFFIRMKSTLTKRGVHIKSSGYKVIHITGRLRLRVSLS
HGRTVPSQIMGLVVVAHALPPPTINEVRIDCHMFVTRVNMDLNIIYCENRISDYMDLTPVDIVGKR
CYHFIHAEDVEGIRHSHLDLLNKGQCVTKYYRWMQKNGGYIWIQSSATIAINAKNANEKNIIWVN
YLLSNPEYKDTPMDIAQLPHLPEKTSESSETSDSESDSKDTSGITEDNENSKSDEKGNQSENSEDPEP
DRKKSGNACDNDMNCNDDGHSSSNPDSRDSDDSFEHSDFENPKAGEDGFGALGPMQIKVERYVE
SESDLRLQNCESLTSDSAKDSDSAGEAGAQASSKHQKRKKRRKRQKGGSASRRRLSSASSPGGLD
AGLVEPPRLLSSPNSASVLKIKTEISEPINFDNDSSIWNYPPNREISRNESPYSMTKPPSSEHFPSPQGG
GGGGGGGGLHVAIPDSVLTPPGADGAAARKTQFGASATAALAPVASDPLSPPLSASPRDKHPGN
GGGGGGGGGAGGGGPSASNSLLYTGDLEALQRLQAGNVVLPLVHRVTGTLAATSTAAQRVYT
TGTIRYAPAEVTLAMQSNLLPNAHAVNFVDVNSPGFGLDPKTPMEMLYHHVHRLNMSGPFGGAV
SAASLTQMPAGNVFTTAEGLFSTLPFPVYSNGIHAAQTLERKED

NPAS3 (153 amino acids) (small) [neuronal PAS3 protein variant 2] Genbank accession AAO17044

MRDFANQGDPPWNLRMEGPPPNTSVKGAQRRRSPSALAIEVFEAHLGSHILQSLDGFVFALNQEG
KFLYISETVSIYLGLSQVELTGSSVFDYVHPGDHVEMAEQLGMKLPPGRGLLSQGTAEDGASSASS
SSQSETPEPVVCFPPASDQFLL

SNX6 (290 amino acids) [nexin sorting protein 6] Genbank accession NP_067072

MTKEEFTKMKQELEAEYLAIFKKTVAMHEVFLCRVAAHPILRRDLNFHVFLEYNQDLSVRGKNK
KEKLEDFFKNMVKSADGVIVSGVKDVDDFFEHERTFLLEYHNRVKDASAKSDRMTRSHKSAADD
YNRIGSSLYALGTQDSTDICKFFLKVSELFDKTRKIEARVSADEDLKLSDLLKYYLRESQAAKDLL
YRRSRSLVDYENANKALDKARAKNKDVLQAETSQQLCCQKFEKISESAKQELIDFKTRRVAAFRK
NLVELAELELKHAKGNLQLLQNCLAVLNGDT

C14ORF11 (233 amino acids) [chromosome 14 open reading frame 11 protein] Genbank accession AAO17041

MNRLPDDYDPYAVEEPSDEEPALSSSEDEVDVLLHGTPDQKRKLIRECLTGESESSSEDEFEKEME
AELNSTMKTMEDKLSSLGTGSSSGNGKVATAPTRYYDDIYFDSDSEDEDRAVQVTKKKKKKQHK
IPTNDELLYDPEKDNRDQAWVDAQRRGYHGLGPQRSRQQQPVPNSDAGLELSCLHDHTLALIAK
GMNHTKTQYRAMFVMNCSINKEEVLRYKASENKEEKAGP

C14ORF10 (453 amino acids) (large) [chromosome 14 open reading frame 10 protein variant 1] Genbank accession AAO17045

MDWKEVLRRRLATPNTCPNKKKSEQELKDEEMDLFTKYYSEWKGGRKNTNEFYKTIPRFYYRLP
AEDEVLLQKLREESRAVFLQRKSRELLDNEELQNLWFLLDKHQTPPMIGEEAMINYENFLKVGEK
AGAKCKQFFTAKVFAKLLHTDSYGRISIMQFFNYVMRKVWLHQTRIGLSLYDVAGQGYLRESDL
ENYILELIPTLPQLDGLEKSFYSFYVCTAVRKFFFFLDPLRTGKIKIQDILACSFLDDLLELRDEELSK
ESQETNWFSAPSALRVYGQYLNLDKDHNGMLSKEELSRYGTATMTNVFLDRVFQECLTYDGEM
DYKTYLDFVLALENRKEPAALQYIFKLLDIENKGYLNVFSLNYFFRAIQELMKIHGQDPVSFQDVK
DEIFDMVKPKDPLKISLQDLINSNQGDTVTTILIDLNGFWTYENREALVANDSENSADLDDT

C14ORF10 (190 amino acids) (small) [chromosome 14 open reading frame 10 protein variant 2] Genbank accession AAO17046

MDWKEVLRRRLATPNTCPNKKKSEQELKDEEMDLFTKYYSEWKGGRKNTNEFYKTIPRFYYRLP
AEDEVLLQKLREESRAVFLQRKSRELLDNEELQNLWFLLDKHQTPPMIGEEAMINYENFLKVGEK
AGAKCKQFFTAKVFAKLLHTDSYGRISIMQFFNYVMRKGDSFLLCLNSLSGNQVGFWSVLL

Appendix 13

PCR primers used in chromosome 14 fine breakpoint mapping of case 10 (Chapter Five)

Marker	Primer sequence (5' → 3')	PCR product size (bp)	$MgCl_2$ (mM)	PCR cycle
[a]B102SP6	F-CTGACTTTAGAAACCAGCTTTCATC R-CAAATGAGAAATTTTGGTCAGAATC	84	1.5	1
[a]B319SP6 (MIPOL1 intron 5)	F-GCCTAGAATTCTCTCTTCTAGATC R-GCCCTGTTTACATGTAAGTTACTG	223	2	1
[a]B368T7 (MIPOL1 intron 3)	F-GTTAGTTGTGAAAAGTTAAGGACC R-ATGTAAAGACTAACAAGTACTACAC	140	2	1
[a]B220SP6	F-TGACACCTCCAGAAGCAAGCAATA R-TATTACCCTAAAATGCCACCTACT	198	2	1
[a]P29SP6 (MIPOL1 intron 11)	F-GCTATGGAAAGAGCATTCCATTAG R-CTGCCTATGTCTTCTGCCTTATC	169	2	1
[a]B305T7 (MIPOL1 intron 11)	F-CACTTCTCTGAACCAGTAAGTTGTC R-AGAGCAGATATTGATCCCAAGGGC	187	2	1
[a]P7T7 (MIPOL1 intron 13)	F-CTACAAAGAAGTTAACTAGCATTC R-AAAGATTATTCCAAGAACTTGCAC	190	2	1
[a]P464T7	F-CTCAGAATCACCGTCATTAATATC R-CACTTTCACCAAATAATTCACAGAG	162	2	1
[a]B132SP6	F-GAAGGTGTCTCCTCATTAAATTATAC R-GCTTCTCATCTGCATAAGTGGG	221	2	2
[a]B156SP6	F-GTTCCTCAGTTGCACTAGCCAC R-TGTGGTGACTTCTACATACACATTC	155	1.5	2
[a]B160SP6	F-CCAGGTTCTGACTAGGGAAAAGG R-CTACGTATTACCAACAACAAGGAAATAG	165	1.5	2
[a]B156T7	F-TATCAGAAAGCAAAGGGGAAGCAG R-GATGGATCCCTCTATGTTCTGGTG	145	1.5	2
[a]B295T7	F-GCTCATGTAAAAACTGATGATTCAC R-TAGGAACTTGCTTAACATAGTTCTC	131	1	1
MPDK1 (MIPOL1 exon 4)	F-TGAACAAGAAACTACGGGGATA R-CTGAAAATTCGTTGATCTCTGC	157	1.5	1
MPDK2 (MIPOL1 exon 5)	F-TTTCTCCTGTTGAAATTCTCAG R-TTGCTCTGGCAATGTCTCTG	194	1.5	1

[a] Primer sequences taken from Matsumoto et al., 1997 (Genomics Vol 45:p11-16)

PCR cycle:

1=94°C-4 mins, 30 cycles of 94°C-30 sec and 58°C-30 sec; followed by 72°C-5 mins final extension

2=94°C-4 mins, 30 cycles of 94°C-30 sec, 58°C-30 sec, 72°C-10 sec; followed by 72°C-5 mins final extension

Appendix 14

PCR primers used for chromosome 4 analysis of case 10 (Chapter Five)

Marker	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**MgCl$_2$ (mM)**	**PCR cycle**
D4S2623	F-AACTAGGCTGCTTCCCAGAT R-GCCAGATACATGGCTAAGGA	201	2	1
D4S2945	F-GGGAGTCATTTCTGGAGGG R-TGAATCTGTAGTTCTGAGCCTTTT	193	1.5	2
D4S2940	F-CTAAGTTGTGCAGCCATGAA R-TGGAACCACTTTTGCAGTAA	219-229	2	3
D4S1651	F-CCCTAACACAAGGGGATTTT R-ACCCTACACCCTTCTTCCAG	200	2	1
D4S1616	F-AGTAATGTGAACAATTCAGTCACTC R-AGCCTCCCAAAGCCAG	246	1.5	1
D4S2842	F-TTGAATTTGCTTTTGGGAGG R-GCTACTGCTTGCTTGGGTTC	132	2	3
D4S2449	F-AAAAGGATCACATAAATCATTTTATGG R-TAGGGAGTTCACATCCAGGC	117	1.5	3
D4S3167	F-ATAAATACCTGTGGTTGGTGCC R-AGTTCATAAGTTTCTGCTCAAATCC	117	1.5	3
D4S1611	F-AGAGTAGTTTCCATCTTTGTTTTC R-GGGCAAGGCTCATCAC	281	1.5	1
D4S1580	F-CGTGGGGGCTATATGATTTG R-ACTCTTTAATATGTTTTGGATCTGG	273	2	1
D4S1573	F-ACATGGAGAATCTTTTAGTAGCA R-CTTTTGAGATACCCCTATCAGT	101-113	2	1
R380DK1	F-GGCAGCACTGGAAGAGGTAA R-CGTTCACGTTAAGCACAGGA	233	1.5	2
R380DK2	F-AGCGGCACTGTCTCATTGAT R-TTTTGGTTGGGAGCTAGAGG	169	1.5	2
R380DK3	F-AGATTGCAGTGAGCCGAGAT R-GGTGACTTCAGCAAGCACAG	180	1.5	2
R380DK4	F-CAGCTTGGGCAACATAGTGA R-GACCTTCAGCCTCAGACAGC	229	1.5	2

PCR cycle:

1=94°C-4 mins, 30 cycles of 94°C-30 sec, 55°C-30 sec; followed by 72°C-5 mins final extension
2=94°C-4 mins, 30 cycles of 94°C-30 sec, 57°C-30 sec; followed by 72°C-5 mins final extension
3=94°C-4 mins, 30 cycles of 94°C-30 sec, 55°C-30 sec, 72°C-30 sec; followed by 72°C-5 mins final extension

Appendix 15

PCR primers of candidate genes tested for rearrangement in case 10 (Chapter Five)

Marker/ Probe name	Gene/ locus	Primer sequence (5' → 3')	PCR product size (bp)	$MgCl_2$ (mM)	PCR cycle
[1]SHGC-83104	ANK2	F-CAGCTGTTTTGCAGAAAATGTGT R-ATAGTAACTTTGGGGGTGGAGGA	346	1.5	1
[1]STS-R39052	ANK2	F-ATCCTTCCAACTTTTCAGAGTTCA R-CCCAGGGAGAGGGAAAGTT	201	1.5	1
AK1	ANK2	F-CCAAGAATCTGGGCGTTTTA R-GTACTCCCAAAGGGCCAAGT	202	1.5	2
AK2	ANK2	F-TCCTCTGAAGGCACAGAGAAA R-TCTGACTGCTCGGTGTCACT	134	1.5	2
TTF1P	NKX2.1	F-ACTACCGCAATCTCCAGTGC R-TCTCTTCCAGTCTCAACGCC	144	1.5	2
[1]SHGC-35528	NKX2.1	F-ACGCGCTTCGACTTTTCTTA R-CCAAAGCTGTTTTATGCCCT	204	1.5	2
N1	NKX2.9	F-TCCGCTGTAAATGCGCTAAT R-CGTCCTGCTCGGGTAAATCT	198	1.5	2
N2	NKX2.9	F-GGTCTCCTGGAACTGGTGAG R-ACGTCCCTCGTGTGTGCT	198	1.5	2
PX1	PITX2	F-CGTGGACTCCTTCGGAACTT R-CCTCCACTTACCGGCTACCG	164	2	3
PX2	PITX2	F-ACTTTCCGTCTCCGGACTTT R-CGCGACGCTCTACTAGTCCT	197	1.5	3
PX3	PITX2	F-CTGACGGGAAAGTGTGTGTG R-CCTCGGAGAGGGAACTGTAA	420	1.5	3
PX4	PTIX2	F-ACTCTATCTCGTCCATGAGC R-CAGGCTCGAGTTACACGTGT	201	2	4
Q9H1	PRO0971	F-CACTGCCTTTGTTCCCTAGC R-CAGCAGCAGTTCCCCATC	204	1.5	2
Q9H2	PRO0971	F-AAATCATGTGGCCCTTTCAA R-GCATTAACACAGGCAATGTGA	219	1.5	2
Q961	T2BP	F-CCCGGGGCTCACACTTAC R-CTCTTCAGCTGTTCGGCTCT	116	1.5	2
Q962	T2BP	F-AGGTCATGAGCCAAGGAATG R-TGGCTTACGACCAACACAAA	198	1.5	2
S1	SLC25A1	F-AGATGTGCAACCGATCCAA R-CCCTTTTAAGTATTTTGCCCCTA	180	1.5	2
S2	SLC25A1	F-TCTTCATGGTGCTGCATAGC R-TCAGAAACCCAGTGGGAAAG	196	1.5	2

PCR cycle:
1=94°C-4 mins, 30 cycles of 94°C-30 sec, 50°C-30 sec; followed by 72°C-5 mins final extension
2=94°C-4 mins, 30 cycles of 94°C-30 sec, 57°C-30 sec; followed by 72°C-5 mins final extension
3=94°C-4 mins, 30 cycles of 94°C-30 sec, 55°C-30 sec, 72°C-30 sec; followed by 72°C-5 mins final extension
4=94°C-4 mins, 30 cycles of 94°C-30 sec, 52°C-30 sec; followed by 72°C-5 mins final extension
[1]Amplimers taken from UniSTS (www.ncbi.nlm.nih.gov)

Appendix 16

PCR primers used for fine breakpoint mapping of cases 11 and 12 (Chapter Five)

Marker	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**MgCl$_2$ (mM)**
1075M22SL	F-ACAAGCCCTTACCCTTTTGG R-GGGCACTCCTTTTTCCTTTC	144	1.5
1075M22SR	F-CCAACAAAGCAACCAAAACA R-TTTTCTGCCTCAGCTCCATT	150	1.5
1078I14SR	F-AACTGGTAATCACCCCACCA R-GCACTGAGCAAAGCTTGACA	152	1.5
66M11SL	F-AATGTGATGCCCTGATGGAT R-TGGGTTTTGATTGACAGCAA	156	1.5
66M11SR	F-TGCAGTCAACAGGTGCTTTC R-CTTTTCCTCCCACTGCTGAG	150	1.5
SDK1	F-TTTCCTGCAGTCCCTGGAT R-GCTTCATAATCGCTGCCAAG	100	1.5
SDK2	F-CTGACAGGCAGCAGTGTCTT R-TCTCCGACTGAGAGGAGGAA	137	1.5
SDK5	F-TTGAAGAGGTCCTTCACATC R-CGGTTATAGCAGCATGATTT	142	1.5
AEX2	F-TGACCAGAGCATTATGGCATT R-AAATTGGTTCAGCAGCAAGG	184	1.5
SDK6	F-GGGCTGTGGAAAAACCATAG R-CTGGTCACCTCCATCTATCACA	100	1.5
SDK7	F-AACTTAAGAGATTTGAGAACTTCA R-TTCACGAACTTTAGGAAACA	98	1.5
I1DK	F-CCTGCTGGACATCAGTGCTA R-TCCCTGTAGAGGCAAGCAGT	232	1.5
I2DK	F-GGCAACAAGAGCAGAACTCC R-GGGCTTCCCTGGTAAAGAAC	208	1.5
I3DK	F-GGATGCAGAAGCAGATAGGC R-GCTCAGTCCACTCCGAACTC	232	1.5
I4DK	F-GTTTGGTAAGCAGGCCACAT R-CCTACCCAGGCCAAACACTA	189	1.5
I5DK	F-GTGAGTGCTCAGGTTGCGTA R-CCATACAGTGAGGCCTGGTT	199	1.5
I6DK	F-AGGCCTCCCTTAACTGTGGT R-CCACCACCCTCTGAAACAGT	225	1.5
I7DK	F-AGGTTTGAGAGGGGTGAGGT R-CATGTCCAAGGGGGATTATG	211	1.5
I8DK	F-CCTGCTGGACATCAGTGCTA R-TCCCTGTAGAGGCAAGCAGT	232	1.5
K1DK	F-GGGCCTTGAGTGAGACTCTG R-CGTTTCCAAAAGCAGAGGAG	128	1.5
K2DK	F-TCCAAGAGCCAACAAAATC R-TGCTGGCCTTAAAAGCCTT	234	1.5
K3DK	F-AGCACTGTTGGGGACATAG R-TTCTGGGTTAGGGGAGGAG	207	1.5
K4DK	F-TCCCCTCTCAAAAGGTGAT R-CATTCTCAGGATGGCTGGT	194	1.5

PCR cycle: 94°C-4 mins, 30 cycles of 94°C-30 sec, 57°C-30 sec; followed by 72°C-5 mins final extension

Appendix 16 continued

PCR primers used for fine breakpoint mapping of cases 11 and 12 (Chapter Five)

Marker	**Primer sequence** (5' → 3')	**PCR product size (bp)**	**MgCl$_2$ (mM)**
D9S1698	F-TAACTGTTCTGAAAAGTGCTCTGG R-GAATTAATGTATGGGTGCTCTTCC	250	1.5
D9S1138	F-GCCTAAGCGGCTGTTTACTG R-AACTTCCTACACCAGATGCACA	174	1.5
D9S159	F-CTTTCTGACGGCAGCCAGGT R-AGCTGGAATGAGTGCTGGGC	293-309	2
D9S260	F-TCATTCTCAATAGCCCACTG R-GCATAACCAGTTCTGCGT	186	2
D9S1928	F-CCTCAGTAAACGAGGAATCC R-GTCTGTGGAGCTGCTCTCTC	155	1.5
D9S1976	F-TTTTTAAATCAAACCTTTGATCATG R-CCCAGACCTAACTGTTGGTAGG	278	1.5
D9S1989	F-TAACTGTTCTGAAAAGTGCTCTGG R-GAATTAATGTATGGGTGCTCTTCC	130	1.5
D9S1881	F-GGCCCAACACCAATGTCAC R-CAATGAGATCATCCATGCAGAGC	220	2
D9S934	F-TTTCCTAGTAGCTCAAGTAAAGAGG R-AGACTTGGACTGAATTACACTGC	224	2
D9S930	F-TGGACAACAGAGTGAGATGC R-TGTAACAAGGGCTATGGGAA	283	2
D9S115	F-CATGAACTTTTACAATGCTTCC R-AGAATAGTCATACCTATACAGG	115-141	1.5
D9S65	F-TGGTCTCTTTTCTGTTGGGG R-GCGGACAATTAGGTTTCAGG	151-161	1.5
D9S752	F-CAGAGGTTGCAGTGAGCTA R-GCAAAGTCAGGCCATTATAC	180	1.5
D9S972	F-CCAGCCAGGTGATGAACTTA R-GCCTGCTGTCGGGGCTGC	125-161	1.5
D9S1014E	F-TCAGGAGTCCCAAAGGTCAG R-GAACAGCAATGAGGGTGCAAG	134	1.5
D9S121	F-CGTAGCGTGGAGTAGCTCT R-GGCAATAGAGTGAGACTCCA	126-142	1.5

PCR cycle: 94°C-4 mins, 30 cycles of 94°C-30 sec, 57°C-30 sec; followed by 72°C-5 mins final extension